THE ANNALS *of* MATHEMATICAL STATISTICS

(FOUNDED BY H. C. CARVER)

THE OFFICIAL JOURNAL OF THE INSTITUTE OF MATHEMATICAL STATISTICS

VOLUME XIV

1943

Math. Econ. Library
HA
1
.A6
Copy 2

THE ANNALS OF MATHEMATICAL STATISTICS

EDITED BY

S. S. WILKS, *Editor*

A. T. CRAIG
W. E. DEMING
T. C. FRY
H. HOTELLING
J. NEYMAN
W. A. SHEWHART

WITH THE COÖPERATION OF

W. G. COCHRAN
J. H. CURTISS
H. F. DODGE
P. S. DWYER
C. EISENHART
W. K. FELLER
P. G. HOEL
W. G. MADOW
A. WALD

The ANNALS OF MATHEMATICAL STATISTICS is published quarterly by the Institute of Mathematical Statistics Mt. Royal & Guilford Aves., Baltimore, Md. Subscriptions, renewals, orders for back numbers and other business communications should be sent to the ANNALS OF MATHEMATICAL STATISTICS, Mt. Royal & Guilford Aves., Baltimore, Md., or to the Secretary of the Institute of Mathematical Statistics, E. G. Olds, Carnegie Institute of Technology, Pittsburgh, Pa. Changes in mailing address which are to become effective for a given issue should be reported to the Secretary on or before the 15th of the month preceding the month of that issue. The months of issue are March, June, September and December. Because of war-time difficulties of publication, issues may often be from two to four weeks late in appearing. *Subscribers are therefore requested to wait at least 30 days after month of issue before making inquiries concerning non-delivery.*

Manuscripts for publication in the ANNALS OF MATHEMATICAL STATISTICS should be sent to S. S. Wilks, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten double-spaced with wide margins, and the original copy should be submitted. Footnotes should be reduced to a minimum and whenever possible replaced by a bibliography at the end of the paper; formulae in footnotes should be avoided. Figures, charts, and diagrams should be drawn on plain white paper or tracing cloth in black India ink twice the size they are to be printed. Authors are requested to keep in mind typographical difficulties of complicated mathematical formulae.

Authors will ordinarily receive only galley proofs. Fifty reprints without covers will be furnished free. Additional reprints and covers furnished at cost.

The subscription price for the ANNALS is $5.00 per year. Single copies $1.50. Back numbers are available at $5.00 per volume, or $1.50 per single issue.

COMPOSED AND PRINTED AT THE
WAVERLY PRESS, INC.
BALTIMORE, MD., U. S. A.

DAW

Entered as second-class matter at the Post Office at Baltimore, Maryland, under the Act of March 3, 1879

the
ore,
om-
Mt.
sti-
gy,
for
the
rch,
ica-
ing.
ssue

TICS
ipts
opy
ever
oot-
on
e to
lties

hout
ost.
.50.
.

379

THE ANNALS *of* MATHEMATICAL STATISTICS

(FOUNDED BY H. C. CARVER)

THE OFFICIAL JOURNAL OF THE INSTITUTE OF MATHEMATICAL STATISTICS

Contents

	PAGE
Some New Methods in Matrix Calculation. HAROLD HOTELLING.	1
On Solutions of the Behrens-Fisher Problem, Based on the *t*-Distribution. HENRY SCHEFFÉ.	35
An Extension of Wilk's Method of Setting Tolerance Limits. ABRAHAM WALD.	45
Asymptotic Formulas for Significance Levels of Certain Distributions. ALFRED M. PEISER.	56
Generalization of Poincaré's Formula in the Theory of Probability. KAI LAI CHUNG.	63
Tables for Testing Randomness of Grouping in a Sequence of Alternatives. FRIEDA S. SWED AND C. EISENHART.	66
Notes:	
A Note on the Best Linear Estimate. A. T. CRAIG.	88
A Note on Tolerance Limits. EDWARD PAULSON.	90
A New Approximation to the Levels of Significance of the Chi-Square Distribution. LEO A. AROIAN.	93
News and Notices.	96
Annual Report of the President of the Institute.	98
Annual Report of the Secretary-Treasurer of the Institute.	100
Constitution and By-Laws of the Institute.	102

Vol. XIV, No. 1 — March, 1943

THE ANNALS OF MATHEMATICAL STATISTICS

EDITED BY

S. S. WILKS, *Editor*

A. T. CRAIG
W. E. DEMING
T. C. FRY
H. HOTELLING
J. NEYMAN
W. A. SHEWHART

WITH THE COÖPERATION OF

W. G. COCHRAN
J. H. CURTISS
H. F. DODGE
P. S. DWYER
C. EISENHART
W. K. FELLER
P. G. HOEL
W. G. MADOW
A. WALD

The ANNALS OF MATHEMATICAL STATISTICS is published quarterly by the Institute of Mathematical Statistics, Mt. Royal & Guilford Aves., Baltimore, Md. Subscriptions, renewals, orders for back numbers and other business communications should be sent to the ANNALS OF MATHEMATICAL STATISTICS, Mt. Royal & Guilford Aves., Baltimore, Md., or to the Secretary of the Institute of Mathematical Statistics, E. G. Olds, Carnegie Institute of Technology, Pittsburgh, Pa. Changes in mailing address which are to become effective for a given issue should be reported to the Secretary on or before the 15th of the month preceding the month of that issue. The months of issue are March, June, September and December. Because of war-time difficulties of publication, issues may often be from two to four weeks late in appearing. *Subscribers are therefore requested to wait at least 30 days after month of issue before making inquiries concerning non-delivery.*

Manuscripts for publication in the ANNALS OF MATHEMATICAL STATISTICS should be sent to S. S. Wilks, Fine Hall, Princeton, New Jersey. Manuscripts should be typewritten double-spaced with wide margins, and the original copy should be submitted. Footnotes should be reduced to a minimum and whenever possible replaced by a bibliography at the end of the paper; formulae in footnotes should be avoided. Figures, charts, and diagrams should be drawn on plain white paper or tracing cloth in black India ink twice the size they are to be printed. Authors are requested to keep in mind typographical difficulties of complicated mathematical formulae.

Authors will ordinarily receive only galley proofs. Fifty reprints without covers will be furnished free. Additional reprints and covers furnished at cost.

The subscription price for the ANNALS is $5.00 per year. Single copies $1.50. Back numbers are available at $5.00 per volume, or $1.50 per single issue.

COMPOSED AND PRINTED AT THE
WAVERLY PRESS, INC.
BALTIMORE, MD., U. S. A.

Entered as second-class matter at the Post Office at Baltimore, Maryland, under the Act of March 3, 1879

By action of the Board of Directors of the Institute, The 1943 Volume of the Annals of Mathematical Statistics is respectfully dedicated to

HENRY LEWIS REITZ

on the occasion of his retirement after twenty-five years of service as Head of the Department of Mathematics of the University of Iowa, and in recognition of his contributions to the initiation and development of mathematical statistics in America.

H. L. Richey

SOME NEW METHODS IN MATRIX CALCULATION[1]

By Harold Hotelling

Columbia University

I. Introduction

1. The increased practical importance of matrix calculation. This paper will be concerned chiefly with methods of finding the inverse of a matrix, and of finding the latent roots and latent vectors, which are also known by a variety of other names associated with particular applications, such as principal axes in geometry and mechanics, and principal components in psychology. These two computational problems are of extremely wide application. The first is closely related to the solution of systems of linear equations, which we shall also consider. In the method of least squares the solution of the normal equations is best carried out with the help of the inverse of the matrix of the coefficients, since at least some of the elements of this inverse matrix are needed in evaluating the results in terms of probability, a vitally necessary step, and since the inverse matrix is useful also in various other ways, such as altering the set of predictors used in a regression equation. Modern statistics also utilizes quadratic and bilinear forms such as the generalized Student ratio [15] for discriminating between samples according to multiple variates instead of one only, the associated discriminant functions [10], the closely related figurative distance of Mahalanobis, Bose and Roy [5] and the critical statistic in an investigation by Wald [28] of the efficient classification of an individual into one of two groups. All these may be calculated very easily from the inverse of a matrix of sums of products, or of covariances or correlations, or from the principal components. Consideration of the relations between two sets of variates [18] may utilize both the inverse of a matrix and a process resembling the calculation of principal components. Similar computational problems arise in applying to sets of numerous variates the contributions to multivariate statistical analysis of R. A. Fisher, S. S. Wilks, W. G. Madow, M. A. Girshick, P. L. Hsu and M. S. Bartlett. Among the non-statistical applications of the inverse matrix and of latent roots and vectors are problems of dynamics, both in astronomy and in airplane design [12], the analysis of stresses and strains in structures [26, 27], and electrical engineering problems [24].

Perhaps no objection to attempts at statistical inference is more common than that the variation of this or that relevant factor has been ignored. For example in dealing with time series the need of allowing for trend and seasonal variation, perhaps by means of a sequence of orthogonal polynomials for trend and of

[1] Revision of a paper presented at the Symposium on Numerical Calculation held Dec. 28, 1941 in New York by the Institute of Mathematical Statistics and the American Statistical Association with the cooperation of the Committee on Addresses in Applied Mathematics of the American Mathematical Society. For the program of the Symposium see the *Annals of Mathematical Statistics* for March, 1942, p. 103.

trigonometric functions for seasonal variation, is well recognized. It is indeed desirable to use regression equations with a liberal number of predictors to eliminate spurious influences, as well as to reduce the error variance, and likewise in other statistical methods. But the computational difficulties in the joint analysis of the desired number of variables have frequently seemed too formidable. We shall see how efficient techniques, in conjunction with efficient machines, can go far to facilitate the use of an appropriate number of variables by reducing the labor to modest dimensions.

While the rise of modern multivariate statistical theory has made available new exact tests of hypotheses in terms of probability over a wide range of cases in which multiple measurements are involved, such measurements have been accumulating on a large scale. In many psychological, anthropometric, astronomical, meteorological and economic fields, actual measurements are available on numbers of variates far greater than have been regarded as amenable, within practical limits, to adequate treatment by the numerical methods generally used. In some instances the number of cases in which complete sets of these variates are available is also large. The 1931 census of India included an extensive sample in which fifty physical variates were measured for each individual. Karl J. Holzinger and his collaborators have worked out and circulated privately a complete matrix of correlations among 78 mental tests. Astronomers have indicated the desirability of a recalculation of the elements of the solar system by means of a gigantic least-square solution with 150 or more unknowns, at the same time deploring the seeming impossibility of this ever being carried out. To apply the methods of modern theoretical statistics to derive from such observations all the important information they contain is an enterprise whose feasibility depends on new numerical methods.

The chief computational problems, apart from those of tabulating and providing convenient approximations for the probability distributions, are (1) the calculation of the many sums of products of pairs of p variates when p is large, and (2) operations on the matrices of these sums of products such as finding the inverse and the principal components. The first problem, which in classical applications of the method of least squares to long series has seemed the heavier, has in a sense been solved by the use of punched cards. A card is used for each case, and all p variates are punched into it. By running the cards repeatedly through a machine wired at each run to select a particular pair of variates, multiply them together, and cumulate the products, this part of the work may be disposed of with great speed. The cost of the machines does at present limit the economical use of this method to rather large numbers, both of variates and of cases. This limit has recently been pushed upward by the introduction of improved multiplying calculators, with high-speed automatic multiplication and squaring locks. But these mechanical advances, in combination with recent discoveries in statistical theory, the increasingly felt need to resort to numerous variates, and the actual existence in many cases of data on such multiple variates, emphasize the need for rapid, economical and accurate calculations with matrices whose elements are sums of products.

Modern machine methods, especially those of the punched-card type, but also those using machines such as the Monroe, Marchant and Fridén, tend to reduce the work of formation of sums of products, in comparison with other operations, to such an extent as to enhance the relative value of methods in which such calculation of direct product-sums is important. Thus products of matrices are much simpler to compute than inverses, and positive than negative powers. Indeed, powers and products of matrices can be computed by means of punched-card machines, and for large matrices this is doubtless the most efficient procedure now available, though considerable rewiring is needed. There is also a possibility, which does not seem too remote, of development of further devices to do this rewiring automatically.

2. Iterative and direct methods. Partitioned matrices. In later sections we shall deal chiefly with certain iterative methods, giving particular attention to the neglected question of limits of error in stopping at any point, and considering the rate of approach to the desired solution. For finding the roots of a matrix and the associated vectors, if the matrix has more than about four rows, it seems clear that an iterative method is the most economical of labor in all but very special cases. On the other hand the problems of solving systems of linear equations and finding the inverse of a matrix do not usually yield readily to iterative methods unless an approximation to the solution is available to begin with. This approximation is not necessarily a very close one, but must not be too wild. It may in some cases be obtained from a general knowledge of the subject.

The Mallock electrical device [22] is capable of solving almost instantaneously ten linear equations in ten unknowns with perhaps two significant digits in each result, though this question of accuracy remains to be elucidated. The combination of this device with the iterative method of Section 7 below, and with the use of partitioning for matrices of more than ten rows, offers what seems at present the best hope for the systematic inversion of large matrices. Since only one of the Mallock machines is in existence (it is in Cambridge, England), some adaptation of the Doolittle or related methods will ordinarily be used. By taking advantage of the possibilities in modern calculating machines of accumulating products to reduce the amount of writing required in the Doolittle method, exceedingly compact and efficient methods have been developed for solving systems of linear equations and for evaluating inverse matrices by Dwyer [7, 8, 9], who utilized the earlier work of Waugh, Kurtz, Horst, Dunlap and Cureton cited by him, and for solving systems of linear equations, by Crout [6]. Dwyer gives valuable bibliographies.

By some of these methods, or from a general knowledge of the subject, one may well obtain approximate solutions correct to a very small number of decimal places, and then by iteration get as many more places as are required, with labor far less than would be necessary to carry through from the beginning the requisite number of places. Further applications of iterative methods arise when a least-square solution is to be revised, either on account of new observa-

tions or because of errors discovered in the original observations or calculations. But however a least-square calculation or the evaluation of any inverse matrix begins, and whatever intermediate steps are taken, it seems advisable to terminate it with the method of Section 7. This combines a check on the previous work, at a labor cost equivalent merely to substituting the values found for the unknowns into the equations, with an improvement in accuracy and a useful limit of error for the unknowns.

In the inversion of large matrices there are important possibilities in the properties of partitioning. For example, a square matrix of $2p$ rows may be partioned into four square matrices a, b, c, d, of p rows, and written

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

If this is multiplied on the right by another partitioned square matrix of $2p$ rows which may be written

$$\begin{bmatrix} A & C \\ B & D \end{bmatrix},$$

where A, B, C, D are square p-rowed matrices, the product

$$\begin{bmatrix} aA + bB & aC + bD \\ cA + dB & cC + dD \end{bmatrix}$$

is identical with the result of partitioning the product of the two original $2p$-rowed matrices. If the second is the inverse of the first, this product is the identical matrix. Consequently, if the first matrix is given, we have for determining its inverse the four matrix equations in A, B, C, D,

$$aA + bB = 1 \qquad aC + bD = 0$$
$$cA + dB = 0 \qquad cC + dD = 1,$$

where 1 stands for the identical matrix of p rows and 0 for the p-rowed matrix consisting entirely of zeros. These equations may be solved just as in elementary algebra except that care must be used to perform matrix multiplications in correct order. Thus

$$A = (a - bd^{-1}c)^{-1}, \qquad B = -d^{-1}cA$$
$$D = (d - ca^{-1}b)^{-1} \qquad C = -a^{-1}bD.$$

These formulae call for inversion of four p-rowed matrices, namely d, $a - bd^{-1}c$, a, and $d - ca^{-1}b$. Without changing the number of such inversions we may choose alternative sets of matrices to invert, with economy of labor in certain cases. For example, if b is easy to invert, we may use for D the expression

$$D = b^{-1}aAbd^{-1}.$$

The formulae and numerical work are further simplified if the given matrix is symmetric. Other modes of partitioning are also possible, and may be valuable in various kinds of numerical work. Another method of obtaining the inverse of a matrix by partitioning is given by Frazer, Duncan and Collar [12, pp. 112–118], who also give an account of general properties of partitioned matrices. In the treatment of relations between two or more sets of variates [18, 31], partitioned matrices appear.

The most efficient method of calculation of a function of a matrix will depend in part on what else is to be calculated. For example, if the latent roots and vectors are needed for any reason as well as the inverse of a matrix, it is better to calculate the former first, and then the determination of the inverse matrix becomes a trivial task; but if the latent roots and vectors are not needed for some other purpose it is usually better not to calculate them but to use a more direct method to obtain the inverse. If in addition to the inverse the determinant is wanted, or many consecutive powers of a matrix, or if a matrix-multiplying machine considerably speedier than present procedures becomes available, a method [3] based on the Cayley-Hamilton theorem that a matrix satisfies its own characteristic equation may be recommended.

Iterative methods have what Whittaker and Robinson [30] call the pleasing characteristic that mistakes do not necessarily spoil the whole calculation, but tend to be corrected at later stages. This of course does not mean that there is no penalty for mistakes. They have an obvious tendency to prolong the number of repetitions required, and if repeated at late stages may actually prevent realization of a substantially correct result. A less obvious consequence of mistakes near the termination of an iterative calculation is that they tend to vitiate any limits of error that may be derived, including those that will be found below. Great care should be used to insure accurate calculation especially in the last stages of any iterative process.

To insure accuracy even before the last stages, and therefore efficiency, a check column consisting of the sums of the elements in the rows of matrices multiplied and added together may well be carried along. In multiplying two matrices only the check column of the second factor is used; it is multiplied by each row of the first factor to obtain the check column for the product. A computer thoroughly experienced with matrix multiplication may dispense with the check column at all stages but the last of an iterative process, relying on the self-correcting property of the process.

A simple but extremely valuable bit of equipment in matrix multiplication consists of two plain cards, with a re-entrant right angle cut out of one or both of them if symmetric matrices are to be multiplied. In getting the element of the ith row and jth column of the product, the ith row of the first factor and the jth column of the second should be marked by a card beside, above, or below it. In writing a symmetric matrix it is convenient to omit the elements below the principal diagonal. The re-entrant right angle is then utilized to mark off the numbers belonging to a particular row.

A report [13] on certain iterative methods of solving linear and other equations and of calculating latent roots and vectors, with engineering applications, was published by R. von Mises and H. Geiringer in 1929. As part of a discussion of certain problems in psychology [16] the present author in 1933 described iterative processes both for solving systems of linear equations and for finding principal components, and later [17] showed how to accelerate convergence to principal components by repeatedly squaring the matrix. Further acceleration of convergence by other devices has been discovered by A. C. Aitken [2]. Dr. Geiringer has also discussed a method of solution of equations involving iteration by small groups of unknowns [14]. The method of Kelley and Salisbury [20] should be noted. It has been used extensively by psychologists. Definite limits of error and measures of rate of convergence for this method are missing. Certain other iterative methods will be discussed in later sections. It will appear that the most-used methods are by no means the best.

Questions regarding the probability of a matrix of covariances satisfying particular conditions of computational significance may in some cases be illuminated with the help of the theory of the variates as a random sample of a larger aggregate. This theory was outlined in the latter part of the paper [16].

II. Linear Equations and Inverse Matrices

3. Accuracy of direct solution of linear equations. The question how many decimal places should be retained in the various stages of a least-square solution and of other calculations involving linear equations has been a puzzling one. It has not generally been realized how rapidly errors resulting from rounding may accumulate in the successive steps of such procedures as, for example, the Doolittle method. In this popular algorism for solving a system of equations

$$\sum_{j=1}^{p} a_{ij}x_j = g_i \qquad (i = 1, \cdots, p),$$

the equivalent of successive eliminations of $x_1, x_2, \cdots, x_{p-1}$ to obtain an equation in x_p alone is accomplished by calculating successively

$$a_{ij\cdot 1} = a_{ij} - a_{i1}a_{1j}/a_{11}, \qquad g_{i\cdot 1} = g_i - a_{i1}g_1/a_{11} \qquad (i, j = 2, 3, \cdots, p),$$

then

$$a_{ij\cdot 12} = a_{ij\cdot 1} - a_{i2\cdot 1}a_{2j\cdot 1}/a_{22\cdot 1},$$

$$g_{i\cdot 12} = g_{i\cdot 1} - a_{i2\cdot 1}g_{2\cdot 1}/a_{22\cdot 1} \qquad (i, j = 3, \cdots, p),$$

and so forth. Let us suppose that each of the a_{ij}'s and g_i's is subject to an error concerning which it is known only that its absolute value does not exceed ϵ. Thus if they are given accurately to k decimal places only, we have $\epsilon = 10^{-k}/2$. Let the actual errors be represented by δa_{ij} and δg_i. If these are small an estimate of the error in $g_{i\cdot 1}$ may be obtained by expanding in a Taylor series and retaining only the linear terms:

$$\delta g_{i\cdot 1} = \delta g_i - \frac{a_{i1}}{a_{11}}\,\delta g_1 - \frac{g_1}{a_{11}}\,\delta a_{i1} + \frac{a_{i1}g_1}{a_{11}^2}\,\delta a_{11}.$$

The closest upper bound for this error obtainable without special assumptions regarding the values of the given quantities is specified by the inequality

$$|\,\delta g_{i\cdot 1}\,| \leq \epsilon\left(1 + \left|\frac{a_{i1}}{a_{11}}\right| + \left|\frac{g_1}{a_{11}}\right| + \left|\frac{a_{i1}g_1}{a_{11}^2}\right|\right).$$

The a's and g's are often correlation coefficients. Any set of normal equations of least squares may be reduced to a form in which this is the case, and this reduction has considerable merits. The various correlation coefficients are frequently of interest in themselves, and their use in the normal equations practically insures that all the quantities appearing at any stage are of the same order of magnitude. This last is a very substantial advantage, partly because of the check column which is customarily carried along, in which each entry is the sum of the other entries in its row. Since the absolute value of a correlation coefficient is less than unity, and since a_{ii} becomes equal to unity, the last inequality gives in this case

$$|\,\delta g_{i\cdot 1}\,| < 4\epsilon,$$

and no closer inequality appears possible. In the same way we find for this case in which the a's are correlation coefficients that

$$|\,\delta a_{ij\cdot 1}\,| < 4\epsilon.$$

Proceeding from these inequalities in the same way, and neglecting the fact that $|\,a_{22\cdot 1}\,| < 1$ though like a_{11} it is put equal to unity in the argument, we find for the errors in $a_{ij\cdot 12}$ and $g_{i\cdot 12}$ the estimated upper bound 16ϵ, with an actual upper bound somewhat higher unless $a_{12} = 0$. Continuing in the same way we find for $a_{ij\cdot 12\cdots(p-1)}$ and $g_{i\cdot 12\cdots(p-1)}$ the estimated limit of error $4^{p-1}\epsilon$, with a possibility of a somewhat higher value up to $4^{p-1}\epsilon/a$, where a is the determinant $|\,a_{ij}\,| < 1$. The rapidity with which this increases with p is a caution against relying on the results of the Doolittle method or other similar elimination methods with any moderate number of decimal places when the number of equations and unknowns is at all large. Thus if $p = 11$ the limit of error exceeds a million times ϵ, indicating that if only one decimal place is wanted in the value of x_p the original correlations must be utilized to at least seven decimals, even if we neglect the additional errors introduced by dropping decimals beyond those retained in the intermediate stages of the calculation. The errors accumulate further during the back solution, so that if all the unknowns are wanted with one-place accuracy it is necessary to use the original correlations with substantially more than seven decimal places. For larger values of p the increase in the error limit is startling. Thus for $p = 27$ (the number of tests reported to be involved in a certain current procedure in classifying military personnel) the limit of error even for the first unknown evaluated is $4^{26}\epsilon$, repre-

senting a loss of about 16 decimal places of accuracy, while the correlations in Holzinger's 78-rowed matrix would need to be carried to no less than 46 places to insure even an approximate accuracy in the first decimal place of one of the regression coefficients in a formula derived by least squares for predicting one of his variates in terms of all the others.

These high limits of error may possibly be reduced in the following ways: (a) a more exact study of the error might be made by means of terms of the Taylor series of orders higher than the first; (b) the positive definite character of a correlation matrix (or other matrix of normal equations) might be utilized in an attempt to arrive at lower limits of error; (c) instead of considering the maximum possible error we might depend on some mutual cancellation of different errors and content ourselves with statements in terms of probability. The compounding of different errors of rounding, which may individually be regarded as having a probability distribution of uniform density over a fixed range, quickly gives rise to an almost exactly normal distribution of known mean and variance, so that the probability approach is attractive. However the limits of error obtained in this way with, for example, a five per cent level of probability of a greater error, though somewhat smaller than the limits associated with certainty, are disappointingly large. Investigations of the types (a) and (b) have not been made; they would apparently be very cumbersome, and (a) might have the effect of increasing the error limits considered above instead of cutting them down. Use of the check column does not provide any safeguard against the errors of rounding appearing in the original correlations, though from the probability standpoint, a carefully devised use of the check column may mitigate the accumulation of errors in successive stages.

To control such errors reliance is often placed in a substitution of the solution obtained in the given equations. This is not completely satisfactory, since under some circumstances large errors in the solution may yield only slight deviations of the left from the right members of the equations, and since some deviations must be expected in any case in which only a limited number of decimals is carried along. Moreover this substitution, even if it reveals the existence of errors, does not usually make clear at once what should be done about them. A recalculation to a larger number of decimal places is horribly laborious. There is here a distinct need of using an iterative process for improving on the solution obtained, and setting definite limits for the errors.

4. The classical iterative method. The iterative method which seems to be the oldest and the most used for solving systems of linear equations, and which may like all other methods of doing this be applied to find the inverse of a matrix, is that of Gauss and Seidel. It seems also to be used in the "method of relaxations" [26], which has been recommended to engineers but lacks limits of error and measures of rate of convergence.

This classical method, starting with any assumed values for the unknowns, begins by changing the value for the first unknown so as to satisfy the first equation; this is possible if the coefficient is different from zero. The revised

set of trial values is then further altered by changing the second unknown so as to satisfy the second equation. Then the third unknown is altered so that the third equation will be satisfied, and so forth. When all the unknowns have been thus altered the cycle may be begun again, and repeated until the differences between consecutive values of each unknown become small enough to indicate a satisfactory convergence. The method converges if the matrix A of the coefficients a_{ij} is positive definite, as it is for the normal equations of least squares, and also in certain engineering applications [7, 8, 9]. Moreover the character of being positive definite insures that each a_{ii} differs from zero, so that the successive adjustments indicated are all actually possible. In the published discussions, proofs of convergence have sometimes been omitted, and in some cases (e.g. [30], Sec. 130) the proofs are incomplete. Even the fuller proofs [13] and [16] fail to give explicit limits for the errors in stopping at any particular stage. But from the discussion [16, pp. 502, 504] it is easy to see that positive numbers d_i and k exist, with $k < 1$, such that the error in the mth estimate of x_i is less than $d_i k^m$. This limit of error diminishes in geometric progression with successive iterations; hence the number of decimal places of accuracy increases approximately in arithmetic progression. The progression is however irregular and the trial values may fluctuate considerably. Numerical determination of limits of error does not appear to be easy. Experience with the method indicates that it is satisfactory only in case a really good approximation is available to begin with, in spite of its universal convergence.

5. An acceleration and extension of the classical iteration. This classical scheme may be improved in the following way if numerous cycles of revision of the trial values are expected to be needed for the requisite accuracy. The first step, consisting of replacing the trial value x_1 by

$$x_1' = (g_1 - a_{12}x_2 - \cdots - a_{1p}x_p)/a_{11}$$

and leaving $x_2, \cdots, x_p$ unchanged, amounts to subjecting the $p + 1$ variables $x_0, x_1, \cdots, x_p$ to the homogeneous transformation

$$\begin{aligned} x_0' &= x_0 \\ x_1' &= (g_1x_0 - a_{12}x_2 - \cdots - a_{1p}x_p)/a_{11} \\ x_2' &= x_2 \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ x_p' &= x_p, \end{aligned}$$

where the symbol x_0, introduced for convenience in order to make these equations homogeneous, is always equal to unity. The matrix of the transformation,

$$T_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ g_1/a_{11} & 0 & -a_{12}/a_{11} & -a_{13}/a_{11} & \cdots & -a_{1p}/a_{11} \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix},$$

is of course singular. If X_0 denote the one-column, $(p+1)$-rowed matrix of the initial trial values, with unity at the head of the column, the column matrix $X_1 = T_1X_0$ is the result of this first operation, again with unity at the head of the column. The trial values obtained by the second operation appear likewise in the column matrix $X_2 = T_2X_1 = T_2T_1X_0$, where

$$T_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & 0 & 0 & \cdots & 0 \\ g_2/a_{22} & -a_{21}/a_{22} & 0 & -a_{23}/a_{22} & -a_{24}/a_{22} & \cdots & -a_{2p}/a_{22} \\ 0 & 0 & 0 & 1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

The result of a complete cycle of substitutions may be written $X_p = T_pT_{p-1}\cdots T_2T_1X_0$, where the matrices T_i are of the same simple character illustrated by T_1 and T_2. This same result will be obtained, because of the associative law of matrix multiplication, if we first calculate numerically the matrix

$$T = T_pT_{p-1}\cdots T_2T_1$$

and then $X_p = TX_0$. (Experience shows that computers need at this point the caution that the matrices must be arranged in their proper order. A good procedure is first to form T_2T_1, then to multiply this by T_3 on the left, etc.). This requires rather more work than the original Gauss-Seidel scheme, and therefore is not worth while if only one cycle of substitutions is needed.

The advantage lies in the fact that T may readily be squared, and T^2X_0 gives a result equivalent to that of two full cycles of iteration by the Gauss-Seidel method. Furthermore, T^2 may be squared to give T^4, which may also be squared, and so on. Obviously k such squarings give a matrix which, when multiplied by X_0, yields the same result as 2^k complete cycles of the original substitutions. In terms of the number k of squarings the number of decimal places of accuracy tends to increase in geometrical instead of arithmetic progression. This modification of the classical method does not seem to have been published heretofore, though both it and the method of Section 7 have been in use by the author and his students since 1936.

R. A. Fisher [11, Sec. 29] has introduced the valuable method of finding the inverse of a matrix A by solving together p systems, each of p equations in p unknowns, with the same matrix A of coefficients, but different columns of unknowns; these several columns of unknowns are the elements of the identical matrix. The technique of carrying this out by any of the methods resembling that of Doolittle is a simple extension involving replacement of the right-hand members of the equations by 1's and 0's and carrying along p such columns instead of one while applying exactly the same linear operations to the rows as in the older problem. This, like the problem of solving linear equations, has been elegantly adapted to efficient calculation with modern machines by Dwyer

[7, 8, 9]. The foregoing iterative methods may also be applied in this case, but the matrix T will be different for the different columns. When the given matrix is symmetric (as is implied by the positive definite character assumed in the proofs of convergence) the number of iterations required is generally cut down because the determination of each column determines also the elements of the corresponding row which lie in other columns. Iteration by groups [14] may well have a place here.

An observation of A. C. Aitken's [1] is noteworthy in connection with the solution of equations with a non-symmetric matrix, and with the finding of the inverse of such a matrix. Writing the equations in the matrix form $AX = G$, we see that the solution $X = A^{-1}G$ is also the solution of the system $(A'A)X = A'G$, where A' is the transverse (also called the transpose or conjugate) of A. Evidently $A'A$ and $A'G$ can be formed by direct multiplications and additions, without divisions. Since $A'A$ is symmetric, any of the methods for solving symmetric equations are applicable to the new system. To find the inverse of A we may first find the inverse of the symmetric matrix $A'A$ and then postmultiply it by A'; for $(A'A)^{-1}A' = A^{-1}$.

6. Roots, norms and convergence of matrices. The *norm* of a matrix A may be defined as the square root of the sum of the products of its elements by their complex conjugates, and denoted by $N(A)$. If A is real and a_{ij} is the element in the ith row and jth column,

$$N(A) = \sqrt{\Sigma\Sigma a_{ij}^2}. \tag{6.1}$$

This is the same function which Wedderburn [29, p. 125] defines as the absolute value of A and denotes by A with a heavy vertical bar on each side. Since it is rather troublesome to avoid confusing this with the determinant of A, we use the notation $N(A)$, though the analogy with the ordinary absolute value of a quantity is very suggestive in connection with proofs of convergence and limits of error obtained by means of the "triangular inequalities" below. Rella [25] gives a different definition of the absolute value of the matrix as the maximum of the absolute values of its roots.

The triangular inequalities, whose proof is easy with the help of the Cauchy inequality, are:

$$N(A + B) \leq N(A) + N(B), \tag{6.2}$$

$$N(AB) \leq N(A)N(B). \tag{6.3}$$

From the last it follows that for any positive integer m,

$$N(A^m) \leq [N(A)]^m. \tag{6.4}$$

Hence if $N(A) < 1$, the limit of $N(A^m)$ as m increases is zero. It then follows that the limit of A^m itself is zero, i.e. that each of its elements approaches zero, because of the definition of the norm.

The identical matrix of p rows, which we shall denote simply by 1, has the

norm $\sqrt{p}$, while a scalar matrix k (i.e. one with the quantity k in each element of the principal diagonal and zeros elsewhere) has the norm $k\sqrt{p}$. The norm of a p-rowed orthogonal matrix is $\sqrt{p}$.

The roots of a square matrix, also known as the latent roots or characteristic roots, are the values $\lambda_1, \cdots, \lambda_p$ of λ for which the determinant obtained by subtracting λ from each element of the principal diagonal vanishes. By expanding this determinant in powers of λ and using a relation between roots and coefficients of an equation, it is evident that the sum of the roots equals the sum of the elements in the principal diagonal. This sum is known as the *trace* of the matrix and denoted by $\mathrm{tr}(A)$. Thus

$$\lambda_1 + \lambda_2 + \cdots + \lambda_p = \mathrm{tr}(A). \tag{6.5}$$

From the definitions of the transverse and norm of A it is plain that

$$[\mathrm{N}(A)]^2 = \mathrm{tr}(AA') \tag{6.6}$$

if A is real.

If $f(x)$ is any polynomial in x, $f(A)$ is a matrix whose roots are known [29, p. 30] to be $f(\lambda_i)$, $(i = 1, 2, \cdots, p)$. In particular, the roots of A^m are λ_i^m. Consequently

$$\lambda_1^m + \lambda_2^m + \cdots + \lambda_p^m = \mathrm{tr}(A^m). \tag{6.7}$$

All the roots of a zero matrix are zero. But the fact that all the roots of a matrix are zero does not necessarily imply that the matrix is zero; for example the roots of

$$\begin{bmatrix} 2 & -1 \\ 4 & -2 \end{bmatrix} \tag{6.8}$$

are both zero. But for real symmetric matrices the vanishing of all the roots does imply the vanishing of the matrix; for the sum of the squares of the elements of a symmetric matrix equals the sum of squares of the roots, since $A = A'$, and by (6.7), (6.6) and (6.1),

$$\Sigma\lambda_i^2 = \mathrm{tr}(A^2) = \mathrm{tr}(AA') = [N(A)]^2 = \Sigma\Sigma a_{ij}^2.$$

Moreover, by continuity considerations, a sequence of p-rowed symmetric matrices must approach zero if all the roots approach zero, and conversely.

From this it is evident that a necessary and sufficient condition that A^m approach zero as m increases, when A is symmetric, is that all the roots of A be less than unity. This provides a sharper criterion of convergence than the requirement that $N(A) < 1$, which is sufficient but not necessary for convergence. The latter is however far easier to apply in most numerical work, since it is far easier to compute $N(A)$ than the greatest root. Moreover it is easy to set an upper bound for $N(A)$ in various ways, of which the crudest is to notice that, by (6.1), $N(A)$ cannot exceed p times the greatest absolute value of any

element of A. Also, the test in terms of the norm is applicable to asymmetric as well as symmetric matrices.

From these considerations regarding the convergence of A^m we deduce at once the following result. If the norm of a square matrix is less than unity, then all the roots are less than unity in absolute value. The converse is not true, as the example (6.8) shows.

For any real square matrix A, symmetric or not,

$$\lambda_1^2 + \lambda_2^2 + \cdots + \lambda_p^2 \leq [N(A)]^2. \tag{6.9}$$

To prove this, we observe first that $2a_{ij}a_{ji} \leq a_{ij}^2 + a_{ji}^2$, and consequently $\mathrm{tr}(A^2) \leq \mathrm{tr}(AA')$. From (6.7) and (6.6) we then have $\Sigma\lambda_i^2 \leq \mathrm{tr}(AA') = [N(A)]^2$. This reasoning shows incidentally that $\Sigma\lambda_i^2$ is real, though the individual roots may be complex.

Not only for investigating convergence, but also in the important but neglected problems of setting definite limits of error after a finite number of steps, the norm is an extremely useful function. If a matrix is to be computed with such accuracy that the error in each element is less than δ, and A is the matrix of errors, the requisite accuracy will according to (6.1) be attained when $N(A) < \delta$. The definition and theorems regarding the norm are valid without any restriction to square matrices, for which alone the roots are defined. For example, we may use the norm to derive an inequality concerning the solution of the system of p linear equations

$$\Sigma a_{ij}x_j = g_i\,,$$

which may be written in matrix form $AX = G$, where A is a square matrix and X and G are matrices each of one column and p rows. From (6.3) we find $N(G) \leq N(A)N(X)$, whence

$$N(X) \geq N(G)/N(A).$$

We shall now deduce a result which seems to be new to matrix theory and which we shall later apply to find limits of error. If A is any matrix such that $1 - A$ is non-singular the identity

$$(1 - A)^{-1} = 1 + A + A^2 + \cdots + A^{m-1} + A^m(1 - A)^{-1}$$

holds, and may be demonstrated exactly as if A were an ordinary scalar quantity. Suppose that $N(A) \leq k < 1$. Taking the norm and using (6.2), (6.3) and (6.4), we have

$$N[(1 - A)^{-1}] \leq p^{1/2} + k + k^2 \cdots + k^{m-1} + k^m N[(1 - A)^{-1}].$$

Since $k < 1$ we may solve for $N[(1 - A)^{-1}]$. Summing the geometric progression, we obtain:

$$N[(1 - A)^{-1}] \leq \frac{p^{1/2} - 1}{1 - k^m} + \frac{1}{1 - k}.$$

This holds for every positive integral value of m, and therefore in the limit when m becomes infinite. Thus we find that

$$N[(1 - A)^{-1}] \leq p^{1/2} - 1 + \frac{1}{1 - k} \tag{6.10}$$

whenever $N(A) \leq k < 1$.

7. An efficient inversion procedure. Let C_0 be an approximation to the inverse of a matrix A, and consider the following sequence of operations. Calculate

$$C_1 = C_0(2 - AC_0), \tag{7.1}$$

and then in turn $C_2, C_3, \cdots$ where

$$C_{m+1} = C_m(2 - AC_m). \tag{7.2}$$

Let us inquire as to the conditions under which the sequence of matrices C_m converges to A^{-1}, the maximum error that may be committed in stopping at any stage, and the rate of convergence. Suppose that C_0 is a good enough approximation to A^{-1} to make the roots of the matrix

$$D = 1 - AC_0 \tag{7.3}$$

all less than unity in absolute value. Then increasing powers of D approach zero, and the convergence of C_m to A^{-1} will follow from the relation

$$C_m = A^{-1}(1 - D^{2^m}), \tag{7.4}$$

which will now be proved by mathematical induction. From (7.1) and (7.3),

$$C_1 = A^{-1}(AC_0)(1 + D) = A^{-1}(1 - D)(1 + D) = A^{-1}(1 - D^2),$$

so that (7.4) is verified for $m = 1$. Now assume (7.4) for a particular value of m, and substitute it in (7.2). This gives

$$C_{m+1} = A^{-1}(1 - D^{2^m})(1 + D^{2^m}) = A^{-1}(1 - D^{2^{m+1}}),$$

which being of the same form as (7.4) completes the induction.

If $N(D) \leq k < 1$ the roots of D are all less than unity in absolute value, as shown in Sec. 6, and the foregoing result holds. Assuming this to be true we now derive an upper bound for the error in C_m in terms of k and $N(C_0)$. According to (7.3),

$$A^{-1} = C_0(1 - D)^{-1}.$$

Hence, by (7.4),

$$C_m - A^{-1} = -A^{-1}D^{2^m} = -C_0(1 - D)^{-1}D^{2^m}.$$

Therefore, by (6.3), (6.4) and (6.10),

$$N(C_m - A^{-1}) \leq N(C_0)k^{2^m}\left(p^{1/2} - 1 + \frac{1}{1 - k}\right). \tag{7.5}$$

This sets an upper bound for the difference between each element of C_m and the corresponding element of A^{-1}. A slightly looser but simpler limit may be obtained from this in terms of the greatest absolute value c of any element of C_0. Since $N(C_0) \leq cp$,

$$N(C_m - A^{-1}) \leq k^{2^m} cp\left(p^{1/2} - 1 + \frac{1}{1-k}\right). \tag{7.6}$$

The great value of this method, whenever a good enough initial approximation is available to make $N(D)$ less than unity, is that the number of decimal places of sure accuracy increases in *geometric* progression, rather than in arithmetic progression as with the usual methods. Consequently this method will always be the most efficient if a sufficiently large number of decimal places is required. Moreover, a limit can be set in advance for the number of iterations that will be required in order to insure any required degree of accuracy. If certainty of correctness in the sth decimal place is required we may choose m so that the right-hand member of (7.6) is less than $10^{-s}/2$. In terms of logarithms to the base 10 the number of decimal places whose accuracy is assured by m iterations is thus at least

$$2^m \mid \log k \mid - \log 2 - \log cp[p^{1/2} - 1 + (1-k)^{-1}]. \tag{7.7}$$

These limits of error can be bettered after some iterations have actually been made. When C_r becomes available we may calculate $k_r = N(1 - AC_r)$, which may be used in place of k in the formulae just derived if m is replaced by $m - r$, and is generally enough smaller than k to make a marked improvement.

The elements of the matrix of errors will actually, of course, be smaller than the norm of this matrix in every practical case, in a ratio fluctuating about p^{-1}. The limits obtained by our formulae can be reached only in case the entire error of the matrix C_m is concentrated in one element, a very unlikely event. Thus the limits given above will usually be quite conservative.

As the iteration proceeds the elements of the matrix $D_m = 1 - AC_m = D^{2^m}$ will diminish rapidly in case of convergence. For this reason it may sometimes be better to calculate C_{m+1} not directly from (7.2), but from the formula

$$C_{m+1} = C_m + C_m D_m \tag{7.8}$$

in which the last term can be regarded as a correction of C_m which will often be very small. This method, however, lacks the self-checking feature, so that its use at the final stage is dubious.

This iterative process has been noticed previously [12, p. 120], but without a limit of error or observation of the geometric progression in the number of accurate digits.

If the initial approximation is not good enough to make $N(D) < 1$, it may be improved by other methods, such as those of Sections 4 and 5, to the point at which this more rapid method becomes applicable. But in some cases (e.g. the second example of §8) the method converges even though $N(D) > 1$, as

may be demonstrated at a later stage at which the norm of the matrix corresponding to D becomes numerically less than unity.

For the mass of least-square and other problems in which the inverse of a matrix is needed, the best procedure appears to begin with one of the methods described by Dwyer [7, 8, 9], carried to a small number of decimal places, and then to calculate D from (7.3), a step equivalent to substituting the approximate solution obtained into the equations. It may then be evident at a glance that the norm of D is so small that the method of the present section will converge rapidly to give as many more places as desired. If $N(D)$ is too large for this, and if gross errors have been eliminated, there is a choice between recalculation from the beginning, the classical iterative process, and the acceleration of this process by matrix-squaring, with perhaps some iteration by small groups. The choice will depend partly on how much the elements of D need to be reduced. The classical iteration (or sometimes the process of this section) is appropriate for correcting a slight excess of $N(D)$ over unity, its matrix-squaring extension for larger alterations.

Let E_0 be the error in C_0, so that $C_0 = A^{-1} + E_0$. Then by (7.1),

$$C_1 = (A^{-1} + E_0)(1 - AE_0) = A^{-1} - E_0 A E_0 .$$

If E_1 is the error in C_1, so that $C_1 = A^{-1} + E_1$, we thus have

$$E_1 = -E_0 A E_0 .$$

If A is symmetric, we naturally take C_0 as a symmetric matrix, and this will cause E_0, C_1, and E_1 also to be symmetric. If also A is positive definite, it will follow from the last equation that E_1 is *negative* definite, or negative semi-definite. Consequently the diagonal elements of C_1 tend to underestimate the corresponding elements of A^{-1}, and never exceed them. Furthermore, the value of a quadratic form whose matrix is A^{-1} will be at least as great as the estimate of it based on C_1. The squares, both of the multiple correlation coefficient and the generalized Student ratio [15], can be expressed as such quadratic forms. Hence both these statistics are slightly underestimated when C_1 is used in place of the true matrix of coefficients. Later approximations C_m do not change the signs of these biases, though they make their magnitudes approach zero in case the conditions for convergence are satisfied, and definite limits converging to zero are easily found for them in such cases from the results above.

8. Illustrations and further comments. We shall indicate symmetric matrices by writing only the elements on and above the principal diagonals.

To illustrate various methods Dwyer [7] has evaluated the inverse of

$$A = \begin{bmatrix} 1.0 & .4 & .5 & .6 \\ & 1.0 & .3 & .4 \\ & & 1.0 & .2 \\ & & & 1.0 \end{bmatrix}$$

as

$$\begin{bmatrix} 2.0710 & -.1913 & -.7759 & -1.0109 \\ & 1.2842 & -.2186 & -.3552 \\ & & 1.3989 & .2732 \\ & & & 1.6940 \end{bmatrix}.$$

If the accuracy of the calculation had been only such as to insure correctness in the first decimal place the approximation to A^{-1} would have been

$$C_0 = \begin{bmatrix} 2.1 & -.2 & -.8 & -1.0 \\ & 1.3 & -.2 & -.4 \\ & & 1.4 & .3 \\ & & & 1.7 \end{bmatrix}.$$

It is easy by mental arithmetic alone, without the use of a machine or side calculations, to see that

$$D = 1 - AC_0 = \begin{bmatrix} -.02 & .02 & 0 & -.01 \\ 0 & 0 & -.02 & .03 \\ .01 & -.01 & 0 & -.02 \\ -.02 & .04 & -.02 & 0 \end{bmatrix}$$

and further that $N(D) = \sqrt{.0052} = .072$. This is so much less than unity that the iteration process of §7 will converge rapidly. As a matter of fact, without determining the sum of the squares of the elements of D we could have observed at a glance that $N(D)$ must be less than four times the greatest absolute value of an element, and thus have a value less than .16. In the same way $N(C_0)$ is seen to be less than 8.4; actually it equals 3.8588. The latter value, with $k = .072$, $p = 4$, substituted in (7.5) gives for the norms of the successive error matrices $E_m = C_m - A^{-1}$,

$$\begin{aligned} N(E_0) &\leq 8.03k &&= .578, \\ N(E_1) &\leq 8.03k^2 &&= .0414, \\ N(E_2) &\leq 8.03k^4 &&= .000216, \\ N(E_3) &\leq 8.03k^8 &&= .000\,000\,0058. \end{aligned}$$

This promises merely that after one application of the iterative process the results will be accurate to one decimal place, which we know already but might not have known for sure in such a case; that a second iteration will give results accurate to three places, and that a third will give results accurate to about eight places. These estimates will however be improved after actually computing C_1. This may well be done by (7.1) if a machine is available; otherwise, and almost as easily, by (7.8) we obtain

$$C_1 = \begin{bmatrix} 2.070 & -.190 & -.776 & -1.011 \\ & 1.282 & -.218 & -.355 \\ & & 1.398 & .274 \\ & & & 1.692 \end{bmatrix}$$

and $N(C_1) = 3.8163$. (We have now passed beyond the stage of easy mental calculation, but might alternatively use the easy upper bound 8.28 for $N(C_1)$, obtained as before.) We shall use this value instead of $N(C_0)$ in (7.5) and at the same time use for k the value of $N(D_1)$, where

$$D_1 = 1 - AC_1 = 1 - AC_0(1 + D) = 1 - (1 - D)(1 + D) = D^2.$$

This is most easily found from D, from which it may be written down directly by mental calculation:

$$D_1 = 10^{-4} \times \begin{bmatrix} 6 & -8 & -2 & 8 \\ -8 & 14 & -6 & 4 \\ 2 & -6 & 6 & -4 \\ 2 & -2 & -8 & 18 \end{bmatrix}$$

The norm of D_1 is seen by the crude method to be less than .0072, and is actually .003212. Taking the latter value for k we have, similarly to (7.5),

$$N(E_m) \leq N(C_1)k^{2^{m-1}} \times 2.00323 = 7.645k^{2^{m-1}}.$$

Thus,

$$\begin{aligned} N(E_1) &\leq .0246, \\ N(E_2) &\leq .000\,0789, \\ N(E_3) &\leq .000\,000\,000\,8. \end{aligned}$$

The reduction in these limits of error is due to the difference between $[N(D)]^2 = .0052$ and $N(D^2) = .003212$.

Using $C_2 = C_1 + C_1D_1$ we obtain:

$$C_2 = \begin{bmatrix} 2.0710366 & -.1912542 & -.7759568 & -1.0109294 \\ & 1.2841486 & -.2185780 & -.3551910 \\ & & 1.3989056 & .2732260 \\ & & & 1.6939852 \end{bmatrix}.$$

From this we calculate

$$D_2 = 1 - AC_2 = 10^{-8} \times \begin{bmatrix} 112 & -164 & -40 & 168 \\ -164 & 288 & -136 & 88 \\ 64 & -128 & 100 & -104 \\ 48 & -32 & -184 & 364 \end{bmatrix},$$

agreeing with the value obtained from the formula $D_2 = D_1^2$, and finally $C_2 = C_2(1 + D_2) =$

$$\begin{bmatrix} 2.071\,038\,458 & -.191\,256\,831 & -.775\,956\,284 & -1.010\,928\,962 \\ & 1.284\,153\,005 & -.218\,579\,235 & -.355\,191\,257 \\ & & 1.398\,907\,104 & .273\,224\,045 \\ & & & 1.693\,989\,071 \end{bmatrix}$$

which as shown above is correct to at least eight decimal places, and doubtless more, in each element. The estimate of A^{-1} obtained by Dwyer by several

direct methods to four places is corroborated by this result excepting for a slight error in the element in his first row and third column.

(ii) Suppose that the approximation in the foregoing example had been even cruder, with determination of the elements of A^{-1} only to the nearest integer. This would give

$$C_0 = \begin{bmatrix} 2 & 0 & -1 & -1 \\ & 1 & 0 & 0 \\ & & 1 & 0 \\ & & & 2 \end{bmatrix},$$

$$D = \begin{bmatrix} .1 & -.4 & .5 & -.2 \\ -.1 & 0 & .1 & -.4 \\ .2 & -.3 & .5 & .1 \\ 0 & -.4 & .4 & -.4 \end{bmatrix}.$$

The sum of the squares of the elements of D is 1.51, so that the norm is greater than unity, and it is not clear at this stage whether the iterative process we have been using will converge or not. But upon computing

$$D^2 = \begin{bmatrix} .15 & -.11 & .18 & .27 \\ .01 & .17 & -.16 & .19 \\ .15 & -.27 & .36 & .09 \\ .12 & .04 & 0 & .36 \end{bmatrix}$$

we find that $N(D^2) = \sqrt{.6093} = .7806$, and since this is less than unity we are assured that the process will converge. We may write immediately, without use of a machine or written side calculation:

$$C_1 = C_0 + C_0D = \begin{bmatrix} 2.0 & -.1 & -.9 & -1.1 \\ & 1.0 & .1 & -.4 \\ & & 1.0 & .3 \\ & & & 1.4 \end{bmatrix}.$$

Utilizing the value of D^2 already determined, we readily find

$$C_2 = C_1 + C_1D^2 = \begin{bmatrix} 2.032 & -.138 & -.848 & -1.056 \\ & 1.138 & -0.42 & -.372 \\ & & 1.182 & .274 \\ & & & 1.558 \end{bmatrix}.$$

From this point on a machine is needed for efficiency. The next step is to calculate D^4, either by squaring D^2 or by the formula $D^4 = 1 - AC_2$; both methods may be used as a check. The result is:

$$D^4 = 10^{-4} \times \begin{bmatrix} 808 & -730 & 1094 & 1330 \\ 20 & 786 & -830 & 890 \\ 846 & -1560 & 1998 & 540 \\ 616 & 80 & 152 & 1696 \end{bmatrix}.$$

We may now consider the accuracy of further approximations, inserting in (7.5) $N(C_2) = \sqrt{13.385572} = 3.659$ in place of $N(C_0)$, $m - 2$ for m, and $k = N(D^4) = .4119$. Thus

$$\begin{aligned} N(E_2) &\leq (9.8807)(.4119) &&= 4.0699 \\ N(E_3) &\leq (9.8807)(.4119)^2 &&= 1.6764 \\ N(E_4) &\leq (9.8807)(.4119)^4 &&= .2844 \\ N(E_5) &\leq (9.8807)(.4119)^8 &&= .00819 \\ N(E_6) &\leq (9.8807)(.4119)^{16} &&= .000\,006\,79. \end{aligned}$$

Because of the roughness of the initial approximation in this case the convergence is rather slow at first, but later it is much accelerated. So far as the limits found above show, five iterations are necessary to be sure of even approximate two-place accuracy in the results (somewhat better limits could be obtained after actually calculating C_2, still better ones from C_3, etc.), but the sixth iteration gives results sure to be accurate nearly to five places. Perhaps the best treatment of a numerical case of this kind is to work out the solution by Dwyer's method to two, three or four places, and then to apply the iterative prócess once, and as many more times as necessary to obtain the required accuracy.

The final step should, for the sake of checking, be a calculation of C_{m+1} from $C_m(2 - AC_m)$, rather than from $C_m + C_m D^{2^m}$.

Upon observing that $N(D) > 1$ we might have used the Seidel process to improve each row of C_0. This process is however extremely slow, and in the present example is markedly inferior to that used above.

(iii) If we start from the result which Dwyer gives to four decimal places as C_0, we obtain

$$D = 1 - AC_0 = 10^{-5} \times \begin{bmatrix} 1 & 4 & -3 & -2 \\ 3 & -2 & 1 & 0 \\ -3 & 3 & -1 & 1 \\ 0 & 2 & 0 & -2 \end{bmatrix}.$$

We find $N(C_0) = 3.8188$, and putting $k = N(D) = .00085$ we have from (7.5),

$$N(E_m) \leq 3.8188\ (.00085)^{2^m}(2.00085) \leq (7.6408)(.00085)^{2^m}.$$

Thus $N(E_1) \leq .000\,0055,$

$N(E_2) \leq .000\,000\,000\,0004.$

9. Certain other methods of successive approximation. A class of methods for solving linear equations, which may be extended to find the inverse of a matrix, is given by Frazer, Duncan and Collar [12, pp. 132–133], generalizing a method of J. Morris. In this method the matrix A of the coefficients in the linear equations, or the matrix to be inverted, is written as the sum of an easily inverted matrix V, for example a diagonal or triangular matrix, and another

matrix W. Then

$$A^{-1} = (1 + V^{-1}W)^{-1}V^{-1} = (1 - f)^{-1}V^{-1},$$

where $f = -V^{-1}W$. If the latent roots of f are all less than unity in absolute value, and *a fortiori* if $N(f) < 1$, the series

$$1 + f + f^2 + f^3 + \cdots = (1 - f)^{-1}$$

converges. To solve the equations $AX = G$, where X and G are column vectors (i.e. matrices of one column) is to determine

$$X = A^{-1}G = (1 - f)^{-1}H,$$

where $H = V^{-1}G$. The method of Frazer, Duncan and Collar is to calculate the successive vectors

$$X_0 = H, \qquad X_1 = H + fX_0, \qquad X_2 = H + fX_1, \cdots, X_r = H + fX_{r-1}, \cdots.$$

It is clear that

$$X_r = (1 + f + f^2 + \cdots + f^r)H.$$

The error in X_r is therefore the vector

$$E_r = f^{r+1}(1 - f)^{-1}H.$$

We may ascertain a limit for the errors if $N(f) \leq k < 1$. Indeed, by (6.3) and (6.10),

$$N(E_r) \leq k^{r+1}\left(p^{1/2} - 1 + \frac{1}{1 - k}\right)N(H),$$

where p is the number of unknowns; and no individual unknown will have an error greater than $N(E_r)$.

Convergence of this method, if existent, may be accelerated by matrix-squaring. Indeed, upon calculating in turn $f^2, f^4, f^8, f^{16}, \cdots$ by repeated squarings, we need only to work with the sequence

$$X_0 = H, \qquad X_1 = (1 + f)X_0, \qquad X_3 = (1 + f^2)X_1,$$

$$X_7 = (1 + f^4)X_3, \qquad X_{15} = (1 + f^8)X_7, \cdots,$$

omitting the intermediate approximations. This will be worth while for solving a single set of equations only in case such great accuracy is required as to demand the use of rather high powers of f. Each squaring of f consists of the formation of p^2 sums of products, so that determination of, say, X_{31} by this method requires $4p^2$ such sums after f has been determined, in addition to the $5p$ involved in finding X_1, X_3, X_7, X_{15}, X_{31} after the squarings. By the method of Frazer, Duncan and Collar the corresponding number of sums of products would be $31p$. Since $4p^2 + 5p < 31p$ only in case $p \leq 6$, it appears that the matrix-squaring is justified only for six or fewer unknowns unless a larger number of terms is required. Furthermore, increasingly high powers of a matrix, to be useful, need usually to be expressed with more and more significant digits.

If more than one system of equations with the same matrix A is to be solved, these methods have the advantage that the same matrix f can be used for all the vectors G of right-hand members. In such cases the value of matrix-squaring is enhanced in comparison with that in which only a single system of equations is to be solved. Determination of A^{-1} is equivalent to solving p such systems in which the several column vectors G together constitute the identical matrix. If more than p of these systems of equations are to be solved it is best to find A^{-1} and then form the various solutions $A^{-1}G$ from the columns G of right-hand members.

It is worth noticing that the matrices $1 + f$, $1 + f^2$, etc., are commutative, as are all rational functions of a single matrix. In difficult cases this may occasionally provide a useful check.

This method differs from the other iterative methods with which we are concerned in that errors of calculation are not automatically corrected by it. This is a serious disadvantage, especially for the inexperienced computer, and makes desirable the careful maintenance of a check column. On the other hand, it does not require any preliminary knowledge of the solution. Indeed, it should be classified rather with the direct than with the iterative procedures on this account.

The critical element in determining the success of this method is the possibility or impossibility of finding suitable matrices V and W, such that V^{-1} can be calculated easily, and such that the elements of $f = -V^{-1}W$ are sufficiently small to make the roots all numerically less than unity. Morris uses for V the matrix derived from A by replacing all the elements above the principal diagonal by zeros. This insures that the corresponding positions in V^{-1} are also occupied by zeros. The other elements of V^{-1} are then determined fairly easily. If the non-diagonal elements of A, which appear in W, are sufficiently small, this fact will insure small enough elements in f to make convergence rapid.

A second method, given by Frazer, Duncan and Collar, chooses for V a diagonal matrix (one having only zero elements except in the principal diagonal), or simply the unit matrix. This choice reduces the labor of inversion to a minimum. Successful convergence will take place when the non-diagonal elements of A are sufficiently small in comparison with those in the diagonal, if V is taken as the diagonal matrix containing the diagonal elements of A.

A third method which may be useful in certain cases, particularly when some but not all of the unknowns are required, is the following. Let A be partitioned:

$$A = \left[\begin{array}{c|c} a & b \\ \hline c & d \end{array}\right],$$

where a and d are square submatrices which, being of lower order than A, are more easily inverted. Let V and W be the correspondingly partitioned matrices

$$V = \left[\begin{array}{c|c} a & 0 \\ \hline 0 & d \end{array}\right], \qquad W = \left[\begin{array}{c|c} 0 & b \\ \hline c & 0 \end{array}\right].$$

Putting $s = a^{-1}b$, $t = d^{-1}c$, we have:

$$f = -\left[\begin{array}{c|c} 0 & s \\ \hline t & 0 \end{array}\right], \qquad f^2 = \left[\begin{array}{c|c} st & 0 \\ \hline 0 & ts \end{array}\right], \qquad f^4 = \left[\begin{array}{c|c} (st)^2 & 0 \\ \hline 0 & (ts)^2 \end{array}\right], \cdots .$$

If only the first q of the p unknowns are required, a and b may be taken as matrices of q rows. If G_1 and H_1 consist respectively of the first q rows of G and H, and if G_2 and H_2 consist of the remaining rows, then $H_1 = a^{-1}G_1$ and $H_2 = d^{-1}G_2$. Then, in case of convergence, the first q rows of the solution are given by

$$(1 + st)[1 + (st)^2][1 + (st)^4] \cdots H_1 - s(1 + ts)[1 + (ts)^2][1 + (ts)^4] \cdots H_2 .$$

Convergence to the correct values is assured here if the norm of any power of st is less than unity, as is true if and only if the absolute values of all the roots of st are less than unity. This is easily seen to be true, since as m increases

$$\lim (ts)^m = t[\lim (st)^{m-1}]s.$$

10. A simple iterative method of solving equations. An entirely different method, whose convergence is independent of the initial trial values, is the following. To solve for the column vector X the equation $AX = G$, we may start with an arbitrary column of trial values X_0 and a scalar constant h, and then for $m = 1, 2, \cdots$ calculate X_m from

$$X_m = hG + (1 - hA)X_{m-1} .$$

If X_m is equal to X_{m-1} it is obviously the desired solution. Otherwise there is an error

$$X_m - X = (hG - X) + (1 - hA)X_{m-1} = (hA - 1)X + (1 - hA)X_{m-1}$$
$$= (1 - hA)(X_{m-1} - X) = \cdots = (1 - hA)^m(X_0 - X).$$

This converges to zero as m increases provided the latent roots of $1 - hA$ are all less than unity in absolute value. If A has only real roots this is equivalent to requiring that they all be between 0 and $2/h$. In particular, if A is a correlation matrix, its roots are all real and positive. Since their sum $= tr(A) = p$, where p is the number of rows, all roots of A lie between 0 and p. Consequently the process will converge in this case if $0 \leq h \leq 2/p$. It is desirable, in order to make the error diminish as fast as possible, to take h as large as is consistent with convergence. In some cases a lower limit than p will be known for the greatest root of A, and then a smaller value than $2/p$ can be taken for h. A limit of error is obviously set by

$$N(X_m - X) \leq [N(1 - hA)]^m N(X_0 - X).$$

This method can of course be applied to find the inverse matrix.

It can also be accelerated by matrix-squaring. If we put $D = 1 - hA$ we have for example,

$$X_8 = (1 + D)(1 + D^2)(1 + D^4)hG + D^8X_0 .$$

The last term will approach zero in case of convergence, and may be omitted in this type of calculation.

Thus accelerated, the method gives decimal places of accuracy increasing in geometric instead of arithmetic progression, and is remarkably simple and straightforward. It is at its best when the roots of A are known to be closely clustered about unity. A criterion of this is that $\Sigma(\lambda_i - \bar{\lambda})^2$ shall have a small value, where $\bar{\lambda}$ is the mean of the p roots λ_i. This sum of squares equals $\Sigma\lambda_i^2 - p$ for a correlation matrix A, and $\Sigma\lambda_i^2 = tr(A^2) = \Sigma\Sigma a_{ij}^2 = p + 2\sum_{i<j} a_{ij}^2$, so that

$$\sum (\lambda_i - \bar{\lambda})^2 = 2\sum_{i<j} a_{ij}^2 .$$

Smallness of this quantity is favorable not only to this iterative method but also to those of §§4 and 5.

11. Use of the characteristic equation for inversion and for finding determinants. A method differing greatly from the others is based on the Cayley-Hamilton theorem that every matrix satisfies its own characteristic equation [29, p. 23; 4, p. 296]. This is the equation

$$f(\lambda) = \begin{vmatrix} a_{11} - \lambda & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} - \lambda & \cdots & a_{2p} \\ \cdots\cdots & \cdots\cdots & \cdots & \cdots\cdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} - \lambda \end{vmatrix}$$

$$= e_p - e_{p-1}\lambda + e_{p-2}\lambda^2 - \cdots + (-)^{p-1}e_1\lambda^{p-1} + (-)^p\lambda^p = 0,$$

where e_r $(r = 1, 2, \cdots, p)$ is the sum of the products r at a time of the roots, and also equals the sum of the r-rowed principal minors of the matrix A. Substituting A for λ, which by the Cayley-Hamilton theorem is legitimate, multiplying by A^{-1}, and transposing a term, yields

$$(11.1) \quad e_pA^{-1} = e_{p-1} - e_{p-2}A + e_{p-3}A^2 - \cdots + (-)^p e_1A^{p-2} + (-)^{p+1}A^{p-1}.$$

This equation provides a direct method of calculating A^{-1} as soon as the elementary symmetric functions e_r of the roots of $f(\lambda) = 0$ have been evaluated. This evaluation may be accomplished by means of Newton's identities [4, p. 243] connecting the elementary symmetric functions with the power-sums. If s_r is the sum of the rth powers of the roots, these formulae give:

$$e_1 = s_1$$
$$e_2 = \tfrac{1}{2}(e_1 s_1 - s_2)$$
$$e_3 = \tfrac{1}{3}(e_2 s_1 - e_1 s_2 + s_3)$$
$$\cdots\cdots\cdots\cdots\cdots$$
$$e_p = \frac{1}{p}(e_{p-1} s_1 - e_{p-2} s_2 + \cdots \pm s_p).$$

The procedure is to calculate in turn A^2, A^3, $\cdots$, A^{p-1}, then to obtain the s's from the diagonals of these matrices, since $s_r = tr(A^r)$, then to obtain the elementary symmetric functions e_1, $\cdots$, e_{p-1} of the roots from Newton's formulae, and to substitute these in the right-hand member of (11.1). It is then only necessary to find and divide by e_p, which equals the determinant of A. For this, and for checking the calculations, there is a choice of methods. We may find the diagonal of A^p, without troubling to compute the whole of this matrix, from the product AA^{p-1} and also, to provide a comprehensive check, from $A^{p-1}A$ or possibly from the product of two powers of A of exponents approximating $p/2$. The sum of these diagonal elements of A^p is s_p, which may be substituted in the last of the Newton formulae above with the quantities previously found to give e_p. An alternative method is to multiply A by its adjoint e_pA^{-1}, which is computed by (11.1), to obtain the determinant e_p.

The total number of multiplications, divisions, and additions is distinctly greater by this method than by efficient direct methods such as that of Dwyer [7, 9]. On the other hand, this method is straightforward and easily checked; the divisions involved are of the simplest character, consisting only of the divisions by 2, 3, $\cdots$, p in Newton's formulae and of the final division of the adjoint matrix by one number; and for large matrices it is ideally adapted for matrix multiplication by means of punched cards. A further very important advantage of this characteristic function method is that it yields considerable additional information as a by-product. Not only the determinant of the matrix but the sums s_r of the principal minors of each order r are determined. Moreover the characteristic equation, whose coefficients would be exceedingly difficult to compute directly from definitions for a large matrix, is by this method made available for the study of the latent roots, which have great interest in themselves for numerous purposes.

The characteristic function method is applicable whether A is symmetric or not. If it is symmetric, the same is true of each of the other matrices appearing in the calculation, so that it is necessary to write only about half the elements.

An illustration using a symmetric matrix has been given by M. D. Bingham [3]. In the illustration below the matrix is not symmetric and has complex double roots and non-linear elementary divisors, so that evaluation of the roots by iterative methods, though possible, would be very slow and laborious, as shown by Aitken [2]. This is indeed the same example used by Aitken in this discussion. But it should be noted that the associated latent vectors, which are determined along with the roots in the iterative processes, require the solution of sets of $p - 1$ linear equations if the roots are found directly by solving the characteristic equation.

$$\text{Let } A = \begin{bmatrix} 15 & 11 & 6 & -9 & -15 \\ 1 & 3 & 9 & -3 & -8 \\ 7 & 6 & 6 & -3 & -11 \\ 7 & 7 & 5 & -3 & -11 \\ 17 & 12 & 5 & -10 & -16 \end{bmatrix}.$$

Then

$$A^2 = \begin{bmatrix} -40 & -9 & 105 & -9 & -40 \\ -76 & -43 & 32 & 44 & 23 \\ -55 & -22 & 62 & 20 & -10 \\ -61 & -25 & 65 & 20 & -7 \\ -40 & -9 & 110 & -14 & -40 \end{bmatrix}, \quad A^3 = \begin{bmatrix} -617 & -380 & 64 & 499 & 256 \\ -260 & -189 & -316 & 355 & 280 \\ -443 & -279 & -106 & 415 & 259 \\ -464 & -300 & -136 & 439 & 292 \\ -617 & -385 & 69 & 499 & 256 \end{bmatrix},$$

$$A^4 = (A^2)^2 = \begin{bmatrix} -1342 & -978 & -2963 & 2444 & 2006 \\ 944 & 522 & -1982 & -10 & 503 \\ -358 & -333 & -2435 & 1307 & 1334 \\ -175 & -243 & -2645 & 1247 & 1355 \\ -1312 & -963 & -2978 & 2444 & 1991 \end{bmatrix} = AA^3 \text{ (check).}$$

From the diagonals of these matrices,

$$s_1 = 5, \quad s_2 = -41, \quad s_3 = -217, \quad s_4 = -17.$$

Calculating the sum of the diagonal elements only of A^5 (on a machine, without listing them separately) from AA^4 and also, as a check, from A^2A^3 we find $s_5 = 3185$. Newton's formulae then give

$$e_1 = 5, \quad e_2 = 33, \quad e_3 = 51, \quad e_4 = 135, \quad e_5 = -225,$$

the last value being that of the determinant of A. We readily find from (11.1),

$$A^{-1} = -\frac{1}{225}\begin{bmatrix} -207 & 64 & -124 & 111 & 171 \\ -315 & 30 & 195 & -180 & 270 \\ -315 & 30 & -30 & 45 & 270 \\ -225 & 75 & -75 & 0 & 225 \\ -414 & 53 & 52 & -3 & 342 \end{bmatrix}.$$

So far, all results by this method are exact, but the division by 225 introduces recurring decimals and therefore a limited validity for the form

$$A^{-1} = \begin{bmatrix} .9200 & -.2844 & .5511 & -.4933 & -.7600 \\ 1.4000 & -.1333 & -.8667 & .8000 & -1.2000 \\ 1.4000 & -.1333 & .1333 & -.2000 & -1.2000 \\ 1.0000 & -.3333 & .3333 & 0 & -1.0000 \\ 1.8400 & -.2356 & -.2311 & .0133 & -1.5200 \end{bmatrix}$$

The characteristic equation

$$f(\lambda) = \lambda^5 - 5\lambda^4 + 33\lambda^3 - 51\lambda^2 + 135\lambda + 225 = 0$$

may in this case be solved readily, since

$$f(\lambda) = (\lambda + 1)(\lambda^2 - 3\lambda + 15)^2.$$

III. Latent Roots and Vectors

12. Direct and iterative methods. If the latent roots but not the latent vectors of a matrix are desired, as for example in a preliminary study of vibra-

tions in machinery being designed, where the important question is whether any root has a positive real part, it is only necessary to find the characteristic equation and to work with it by the methods of the theory of equations. The coefficients in the characteristic equation are the sums of the r-rowed principal minors ($r = 1, 2, \cdots, p$), and are expeditiously found directly from this definition for matrices of four or fewer rows. For large matrices, however, the calculation of so many large overlapping determinants is wasteful of effort, since many virtually equivalent calculations must be done repeatedly. Indeed, calculation by determinants in a great many situations, including the solution of linear equations, is open to this objection. The methods of §11 yield the characteristic function in a manner which, for large matrices, appears to be the best available, excepting perhaps the new method of Samuelson [25a].

When, as is commonly the case, the latent vectors are desired, a straightforward calculation directly from the definitions would require not only setting up and solving the characteristic equation, but also the solution, in the case of each root, of the set of linear equations in p unknowns whose matrix is obtained from the characteristic matrix by substituting the particular root for λ. It is this solution of linear equations that aggravates greatly the computational labor when direct methods are used.

An ingenious method has been used by R. A. Fisher [11, pp. 299 ff.]. Starting with a four-rowed determinant whose elements are linear functions of an unknown θ, Fisher calculates the value of the determinant for selected values of θ, and then by interpolation using divided differences finds the largest value of θ making the determinant zero. The point of the divided difference method is that it avoids the direct calculation of the determinant for more than a few values of θ, replacing it essentially by calculation of the fourth-degree polynomial in θ from its differences and using the fact that the fourth divided differences are constant. The linear equations are then solved in a direct manner. If applied to large matrices this would be very laborious, but it compares favorably with calculation directly from definitions in the manner suggested by reading books on algebra and solid analytic geometry. But even with large matrices Fisher's method may perhaps be the best in certain cases, e.g. if all that is desired is the root of median absolute value and if this root is real, or if it is desired to find a few real roots that are close together, with numerous others greater and another numerous group less than these. This is because the iterative methods give the real roots in the order of their absolute values, beginning with the greatest, but with the possibility of obtaining them in the opposite order by first inverting the matrix. The Mallock electrical device [22] may be used to calculate determinants, and thus to apply this method.

If A and B are p-rowed matrices and B is non-singular, the determinantal equation $|A - \lambda B| = 0$ is equivalent to $|AB^{-1} - \lambda| = 0$ and also to $|B^{-1}A - \lambda| = 0$. The column vectors X_i satisfying $(A - \lambda B)X = 0$ also satisfy $(B^{-1}A - \lambda)X = 0$ and the row vectors V_i satisfying $V_i(A - \lambda_i B) = 0$ also satisfy $V_i(AB^{-1} - \lambda_i) = 0$. If A and B are symmetric, $V_i = X_i'$. Thus

any problem of this type is reducible to that of finding latent roots and vectors, upon calculating B^{-1} by any method and multiplying in either order by A.

The fundamental iterative method for finding latent roots and vectors of A begins with an arbitrary matrix X_0 of a single column. This column vector is premultiplied by A to obtain a new column vector X_1. If, as is possible though unlikely, the elements of X_1 are proportional to those of X_0, they constitute one of the latent vectors of A, and the factor of proportionality is the corresponding root, for then X_0 and X_1 are solutions of the matrix equation $(A - \lambda)X = 0$. It should be observed that the latent vector is determined only to within an arbitrary scalar factor of proportionality, though we may sometimes find it convenient to normalize the vector by choosing the factor in such a way that the sum of the squares of the elements, which equals the square of the norm, is unity.

If X_1 is not proportional to X_0, the operation may be repeated by calculating $X_2 = AX_1$, then $X_3 = AX_2$, and so on. If these vectors are then normalized, or if they are divided by, say, their respective first elements, then the other elements will (in the cases of greatest practical importance) gradually approach stable values which will determine one of the latent vectors, while the successive factors of proportionality will approach the corresponding root. The convergence of this process is however apt to be rather slow. Fortunately there are several known ways of accelerating it.

Matrix-squaring is the first of these methods of accelerating convergence [17, 19]. It is clear that $X_t = A^t X_0$. Consequently one application of the iterative process with A^t is equivalent to t iterations with A. It is relatively easy to square A, and then by repeated squarings to form A^4, A^8, A^{16}, etc. The economical limit of this process is determined partly by the necessity of retaining more and more digits in the successively higher powers, but up to a point not yet determined exactly it presents very great advantages. For proceeding to the determination of latent roots of other than the maximum absolute value, with their associated vectors, this method lends itself to further shortcuts [17, 2], which seem to give it an advantage over an older method [13].

Another method of accelerating convergence, introduced by A. C. Aitken, and referred to by him as the δ^2-method, uses the ratio $\phi(t)$ of an element of X_{t+1} to the corresponding element of X_t in the function

$$\frac{\phi(t+1)\phi(t-1) - [\phi(t)]^2}{\phi(t+1) - 2\phi(t) + \phi(t-1)},$$

which converges rapidly toward the root λ_1 of greatest absolute value. If a constant c is subtracted from all three of the quantities $\phi(t+1)$, $\phi(t)$ and $\phi(t-1)$ before computing the foregoing function the result is unchanged. This fact reduces greatly the computational labor, since the decimal places of λ_1 already determined are common to all three.

If A is symmetric and we form the scalar products of $X_t = A^t X_0$ with itself

and with X_{t+1} we have

$$X_t'X_t = X_0'A^{2t}X_0\,, \qquad X_{t+1}'X_t = X_0'A^{2t+1}X_0\,.$$

The ratio of these two scalars gives an estimate of λ_1 which on the basis of the ratios of consecutive elements in a given place in the trial vectors would not be reached until a later stage of convergence, corresponding in fact to twice as many iterations. Aitken has pointed out the great value of this procedure for finding the root (but not the latent vector), and has extended the idea to asymmetric matrices, where there is a complication because of the existence of two latent vectors for each root, one determined by premultiplying by A, the other by A'.

The comprehensive paper [2] of Aitken gives an extremely valuable account of the whole problem and processes of finding the latent roots and vectors, including a survey of the various cases arising when there are multiple roots, complex roots, and non-linear elementary divisors. This paper should be studied carefully by anyone with any substantial numerical problem of this kind.

A method using rotations of two variables at a time has been devised by T. L. Kelley [21].

The remainder of this paper will be concerned with some results, believed to be new, by which useful upper limits can be set for the errors of the results yielded by iteration for latent roots and vectors of a symmetric matrix. To find such limits of error for asymmetric matrices appears to be a much more difficult and as yet unsolved problem.

13. Accuracy of iteration with symmetric matrices. If A is symmetric, as it is in most statistical problems (though with some exceptions, as in [18]), the roots are all real and the elementary divisors are linear. Moreover there exist an orthogonal matrix H and a diagonal matrix

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & 0 & \cdots \\ 0 & \lambda_2 & 0 & \cdots \\ 0 & 0 & \lambda_3 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{bmatrix}$$

such that

$$A = H\Lambda H'. \tag{13.1}$$

Since H is orthogonal, $HH' = 1$, and therefore

$$\Lambda = H'AH, \qquad A^t = H\Lambda^t H'. \tag{13.2}$$

We may associate with the successive trial vectors $X_t = AX_{t-1} = A^tX_0$ the vectors $Y_t = H'X_t$; then $X_t = HY_t$. From these equations and the second of (13.2) it is clear that

$$Y_t = H'X_t = H'A^tX_0 = \Lambda^tH'X_0 = \Lambda^tY_0\,.$$

Hence, if the elements of Y_0 are $y_1, \cdots, y_p$,

$$Y_t = \begin{bmatrix} y_1\lambda_1{}^t \\ y_2\lambda_2 \\ \cdot \\ \cdot \\ \cdot \\ y_p\lambda_p{}^t \end{bmatrix}.$$

Now let α_{kt} be the scalar product of X_t and X_{t-k}:

$$\alpha_{kt} = X_t'X_{t-k} = Y_t'Y_{t-k} = \Sigma y_i^2\lambda_i^{2t-k}; \tag{13.3}$$

and let

$$\nu_{kt} = \frac{\alpha_{kt}}{\alpha_{0t}} = \frac{\Sigma y_i^2\lambda_i^{2t-k}}{\Sigma y_i^2\lambda_i^{2t}}. \tag{13.4}$$

If A has a negative root this fact will become evident after a certain stage in the iteration used to obtain this root by an alternation of sign of the numbers in any one position in consecutive trial vectors. However A^2, which as pointed out in §12 may well be calculated anyhow, has only positive roots, which are the squares of the roots of A, and has the same latent vectors as A. Hence we shall have results of sufficient generality for real symmetric matrices if we assume that all roots of the matrix with which we work are positive or zero, i.e. that it is positive definite or semi-definite. Let us choose the notation so that

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0.$$

Then if $k \geq 0$,

$$\alpha_{0t} = \Sigma y_i^2\lambda_i^{2t} \leq \lambda_1^k\Sigma y_i^2\lambda_i^{2t-k} = \lambda_1^k\alpha_{kt}.$$

Hence, by (13.4),

$$\lambda_1 \geq [\nu_{kt}]^{-1/k}. \tag{13.5}$$

It is known [23] that if $a_1, \cdots, a_p, c_1, \cdots, c_p$ are any positive numbers, the function

$$\left(\frac{c_1a_1^k + \cdots + c_pa_p^k}{c_1 + \cdots + c_p}\right)^{1/k}$$

increases monotonically with k. Putting $c_i = y_i^2\lambda_i^{2t}$, $a_i = \lambda^{-k}$ if $\lambda_i \neq 0$, and $c_i = a_i = 0$ if $\lambda_i = 0$, we find that the right-hand member of (13.5) decreases monotonically as k increases. Hence the best of these lower bounds for λ_1 is that corresponding to the least value of k that can be used, namely $k = 1$. Consequently the lower bound to be recommended for λ_1 is given by

$$\lambda_1 \geq \frac{1}{\nu_{1t}}. \tag{13.6}$$

From (13.4) it is easily seen that this lower bound approaches λ_1 when t increases, provided $y_1 \neq 0$.

An upper bound for λ_1 is available from the fact that the sum of the tth powers of the roots is the trace of A^t. Since we assume all $\lambda_i \geq 0$ this gives

$$\lambda_1 \leq (\text{tr } A^t)^{1/t}.$$

That this upper limit converges to λ_1 when t increases is easily seen from (6.7) upon consideration of $\log (\Sigma\lambda_i^t)^{1/t}$.

A lower limit alternative to that of (13.6) is also available from tr (A^t), and likewise converges to λ_1. Indeed, since λ_1 is the greatest root, we have

$$\lambda_1 \geq (\text{tr } A^t/p)^{1/t}.$$

We now seek limits of accuracy for the latent vector corresponding to λ_1 and estimated by X_t. If we call this vector X, and define $Y = H^{-1}X$, then $\lim Y_t^* = Y^*$, where Y^* is the normalized form of Y. Y^* has as its ith element

$$\lim_{t\to\infty} \frac{y_i\lambda_i^t}{\sqrt{\Sigma y_j^2\lambda_j^{2t}}}.$$

If $y_1 \neq 0$, and $\lambda_1 > \lambda_2 \geq \lambda_3$, this limit is ± 1 if $i = 1$, and is otherwise 0. We take the value of Y to be

$$\begin{bmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix}$$

in this case. If λ_1 is a multiple root, the limit of X_t^* will depend on the initial values y_i.

A useful measure of the closeness of approach of X_t to X is the "correlation coefficient" $r_t = X^{*\prime}X_t^* = Y^{*\prime}Y_t^* = \dfrac{y_1\lambda_1^t}{\sqrt{\Sigma y_j\lambda_j^{2t}}}$,

which obviously approaches unity as t increases if $y_1 \neq 0$ and λ_1 is a *simple* root, or if $\lambda_1 = \lambda_2 = \cdots = \lambda_s > \lambda_{s+1}$ and we arrange our definitions so that $y_1 \neq 0$ and $y_2 = \cdots = y_s = 0$. In terms of the notation previously introduced, $r_t = \dfrac{y_1\lambda_1^t}{\sqrt{\alpha_{ot}}}$. The sum of the squares of the differences of corresponding elements of the normalized vectors X^* and X_t, i.e. $[N(X^* - X_t^*)]^2$, is $2(1 - r_t)$, and therefore approaches zero as r_t approaches unity. We shall seek for r_t a lower limit approaching unity as t increases.

Let us now put

$$w_{it} = \frac{y_i^2\lambda_i^{2t}}{\Sigma y_j^2\lambda_j^{2t}} = \frac{y_i^2\lambda_i^{2t}}{\alpha_{ot}}. \qquad \text{Then } r_t^2 = w_{1t} \text{ and } \sum_{i=1}^{p} w_{it} = 1.$$

For $k \geq 1$,

$$\alpha_{kt} = \Sigma y_i^2 \lambda_i^{2t-k} \geq y_2^2 \lambda_2^{2t-k} + \cdots + y_p^2 \lambda_p^{2t-k} \geq \lambda_1^{-k}(y_2^2 \lambda_2^{2t} + \cdots + y_p^2 \lambda_p^{2t})$$
$$= \lambda_1^{-k} \alpha_{ot}(w_{2t} + \cdots + w_{pt}) = \lambda_1^{-k} \alpha_{ot}(1 - w_{1t}) = \lambda_1^{-k} \alpha_{ot}(1 - r_t^2).$$

$$\therefore r_t^2 \geq 1 - \frac{\lambda_1^k \alpha_{kt}}{\alpha_{ot}} = 1 - \nu_{kt} \lambda_1^k.$$

This unfortunately is not a very useful lower bound for r_t, since it approaches *zero*, not unity, as t increases.

A more satisfactory result is obtained as follows. Let $\eta_i = \lambda_i^{-1}$. Then $\nu_{kt} = \sum_{k=1}^{p} w_{it} \eta_i^k$. For any value of t we may consider a distribution of a variate taking the positive values $\eta_1, \cdots, \eta_p$ with the positive weights, or probabilities, w_{it}. The kth moment of this distribution about 0 is ν_{kt}. In particular the first moment is ν_{1t}, and is evidently at least equal to η_1, which is the least of the η_i. The standard deviation is $\sigma = \sqrt{\nu_{2t} - \nu_{1t}^2}$. As t increases, ν_{1t} will approach η_1 and σ will approach zero. Hence, if $\lambda_1 > \lambda_2$, a stage will eventually be reached at which $\nu_{1t} < \eta_2$. Let

$$k = \frac{\eta_2 - \nu_{1t}}{\sigma}.$$

By the Tchebychef-Bienaymé inequality,

$$w_{2t} + \cdots + w_{pt} \leq \frac{1}{k^2},$$

and therefore

$$r_t^2 = w_{1t} \geq 1 - \frac{\nu_{2t} - \nu_{1t}^2}{(\eta_2 - \nu_{1t})^2},$$

provided t is large enough so that $\nu_{1t} < \eta_2$. This lower bound approaches unity, as desired, when t increases.

If $\lambda_1 = \lambda_2, \cdots = \lambda_k > \lambda_{k+1}$, the same proof shows that

$$w_1^{(t)} + \cdots + w_k^{(t)} \geq 1 - \frac{\nu_{2t} - \nu_{1t}^2}{(\eta_{r+1} - \nu_{1t})^2},$$

provided $\nu_{1t} < \eta_{r+1}$. The left member is the correlation of X_t with that one of the k-parameter family of latent vectors corresponding to the multiple root for which the correlation is a maximum.

In order to utilize these results we need a lower bound for η_2, or for η_{r+1}. In case $\lambda_1 \neq \lambda_2$ this requires an upper bound for λ_2. Such an upper bound may be found at the next stage through working with the reduced or "deflated" matrix used in [17]. This is $A_1 = A - \lambda_1 XX'$, where X is the normalized latent vector corresponding to λ_1; and $\lambda_2^t \leq \text{tr } (A_1^t)$. .

Since we have arrived at a definite lower limit for r_t which approaches unity

as the iterative process proceeds, and since we have found for λ_1 upper and lower bounds converging to it, a solution has been found for the troublesome problem of the degree of accuracy in stopping at any stage of the iteration for finding the greatest root and the associated latent vector. It would be possible to go on to find from these results appropriate inequalities for A_1, and then by repetition of the above arguments, for λ_2 and the second latent vector; and then likewise for the second reduced matrix A_2 and the further roots, vectors, and reduced matrices in this cyclic order. These steps may well be taken by the computer who has mastered the above argument in connection with a numerical example.

BIBLIOGRAPHY

[1] A. C. Aitken, "Studies in practical mathematics. I. The evaluation, with applications, of a certain triple matrix product," *Proc. Roy. Soc. Edinburgh*, Vol. 57 (1937), pp. 172–181.

[2] A. C. Aitken, "Studies in practical mathematics, II. The evaluation of the latent roots and latent vectors of a matrix," *Proc. Roy. Soc. Edinburgh*, Vol. 57 (1937), pp. 269–304.

[3] M. D. Bingham, "A new method for obtaining the inverse matrix," *Jour. Amer. Stat. Assoc.*, Vol. 36 (1941), pp. 530–534.

[4] M. Bôcher, *Introduction to Higher Algebra*, New York, 1907, 321 pp.

[5] R. C. Bose and S. N. Roy, "The distribution of the Studentized D^2-statistic," *Sankhya*, Vol. 4 (1938), pp. 19–38.

[6] Prescott D. Crout "A short method for evaluating determinants and solving systems of linear equations with real or complex coefficients." *Trans. Am. Inst. Elec. Engineers*, Vol. 60 (1941).

[7] Paul S. Dwyer, "The solution of simultaneous equations," *Psychometrika*, Vol. 6 (1941), pp. 101–129.

[8] Paul S. Dwyer, "The Doolittle technique," *Annals of Math. Stat.*, Vol. 12 (1941), pp. 449–458.

[9] Paul S. Dwyer, "Recent developments in correlation technique," *Jour. Amer. Stat. Assoc.*, Vol. 37 (1942), pp. 441–460.

[10] R. A. Fisher, "The use of multiple measurements in taxonomic problems," *Annals of Eugenics*," Vol. 7 (1936), pp. 179–188.

[11] R. A. Fisher, "*Statistical Methods for Research Workers*, London, 7th ed., 1938, 256 pp.

[12] R. A. Frazer, W. J. Duncan and A. R. Collar, *Elementary matrices and some applications to dynamics and differential equations*, Cambridge Univ. Press, 1938. 416 pp.

[13] R. von Mises and Hilda Pollaczek-Geiringer, "Zusammenfassende Berichte. Praktische Verfahren der Gleichungsauflösung." *Zeitschrift für angewandte Math. und Mechanik*, Vol. 9 (1929), pp. 58–77 and 152–164.

[14] Hilda Geiringer, "On the numerical solution of linear problems by group iteration," *Bull. Amer. Math. Soc.*, Vol. 48 (1942), p. 370.

[15] Harold Hotelling, "The generalization of Student's ratio," *Annals of Math. Stat.*, Vol. 2 (1931), pp. 360–378.

[16] Harold Hotelling, "Analysis of a complex of statistical variables into principal components," *Jour. Educ. Psych.*, Vol. 24 (1933), pp. 417–441 and 498–520. (Reprints available only from publishers, Warwick and York, Baltimore.)

[17] Harold Hotelling, "Simplified calculation of principal components," *Psychometrika*, Vol. 1 (1936), pp. 27–35.

[18] Harold Hotelling, "Relations between two sets of variates," *Biometrika*, Vol. 28 (1936), pp. 321–377.

[19] R. C. J. Howland, "Note on a type of determinantal equation," *Phil. Mag.*, Series 7, Vol. 6 (1928), pp. 839–842.

[20] Truman L. Kelley and Frank S. Salisbury, "An iteration method for determining multiple correlation constants," *Jour. Amer. Stat. Assn.*, Vol. 21 (1926), pp. 282–292.

[21] Truman L. Kelley, *Essential traits of mental life*, Harvard, 1935, 145 pp. Chap. 1.

[22] *The Mallock electrical calculating machine.* Reprint from *Engineering* (London), June 22, 1934. 8 pp.

[23] Nilan Norris, "Inequalities among averages," *Annals of Math. Stat.*, Vol. 6 (1935), pp. 27–29.

[24] Louis A. Pipes, "Matrix theory of oscillatory networks," *Jour. Applied Physics*, Vol. 10 (1939), pp. 849–860.

[25] T. Rella, "*Über den absoluten Betrag von Matrizen*," International Congress of Mathematicians at Oslo, 1936 (derives λ_1 as the absolute value from 5 postulates, which are proved independent).

[25a] P. A. Samuelson, "A method of determining explicitly the coefficients of the characteristic equation," *Annals of Math. Stat.*, Vol. 13 (1942), pp. 424–429.

[26] R. V. Southwell, *Relaxation methods in engineering science, a treatise on approximate computation*, Oxford Univ. Press, 1940. 252 pp.

[27] G. Temple, "The general theory of relaxation methods applied to linear systems," *Proc. Roy. Soc. London*, Vol. 169 A (1939), pp. 476–500.

[28] A. Wald, "The classification of an individual into one of two groups." (Unpublished).

[29] J. H. M. Wedderburn, *Lectures on matrices*, New York, 1934. 200 pp.

[30] E. T. Whittaker and G. Robinson, *The calculus of observations*, London, 1924. 395 pp.

[31] S. S. Wilks, "Certain generalizations in the analysis of variance," *Biometrika*, Vol. 24 (1932), pp. 471–494.

Of interest in relation to the subject of this paper, though not mentioned in the text, are the following:

Arnold Dresden, "On the iteration of linear homogeneous transformations," *Bull. Amer. Math. Soc.*, Vol. 48 (1942), pp. 577–579 and 949.

Paul Horst, "A method for determining the coefficients of a characteristic equation," *Annals of Math. Stat.*, Vol. 6 (1935), pp. 83–84.

A. T. Lonseth, "Systems of linear equations with coefficients subject to error," *Annals of Math. Stat.*, Vol. 13 (1942), pp. 332–337.

Rufus Oldenburger. "Infinite powers of matrices and characteristic roots," *Duke Math. Jour.*, Vol. 6 (1940), pp. 357–361.

ON SOLUTIONS OF THE BEHRENS-FISHER PROBLEM, BASED ON THE *t*-DISTRIBUTION

By Henry Scheffé

Princeton University

1. The Problem. The problem [1, 2] is the interval estimation[1] of the difference of the means of two normal populations when the ratio of the variances of the populations is unknown. The reader who wishes to see the present solution before considering theoretical details will find it recapitulated in the Summary at the end and will want to refer to the following notation: $(x_1, x_2, \cdots, x_m)$ and $(y_1, y_2, \cdots, y_n)$ are random samples from normal populations with means α and β, and variances μ and ν, respectively. Define $\delta = \alpha - \beta$. We assume $m \leqq n$, and that the variates in each sample are in the order of observation, or else have been randomized.

Recently Neyman [3] has called attention to a solution which we shall designate as (B), and which is a special case of an unpublished solution of Bartlett[2]. It will be simpler to describe (B) later, but we mention now that it has the following advantages: (*i*) its validity does not depend on the values of unknown parameters, (*ii*) the required computations are simple, and (*iii*) only existing tables are needed,——the widely available Fisher *t*-tables. An unsatisfactory aspect of (B) is that when the sample sizes are unequal, $n - m$ of the variates y_i are completely discarded. The solution below shares with (B) the advantages (*i*), (*ii*), (*iii*); indeed, it is identical with (B) when $n = m$, but when $n \neq m$ it is free from the above objection.

2. Simple Solution. We begin with a simple restricted approach; later we will review the result from a somewhat broader standpoint. If random variables $d_1, d_2, \cdots, d_m$ are independently normally distributed with mean δ and variance σ^2, and if L and Q are defined from

$$L = \sum_{i=1}^{m} d_i/m, \qquad Q = \sum_{i=1}^{m} (d_i - L)^2,$$

then $m^{\frac{1}{2}}(L - \delta)/\sigma$ and Q/σ^2 are independently distributed; the former is a normal variable with zero mean and unit variance; the latter, $Q/\sigma^2 = \chi^2_{m-1}$, where χ^2_k is a generic notation for a random variable distributed according to the χ^2-law with k degrees of freedom. The quotient

---

[1] We treat the problem from the standpoint of confidence intervals, rather than significance tests, since when the former are available for δ so is a whole class of the latter, namely for any hypothesis $\delta = \delta_0$, for all δ_0. Furthermore, questions of the existence of "best" tests and "best" confidence intervals are closely related [5a].

[2] How far Bartlett followed the path of this paper is not clear from the brief mention of his results by Welch [4], except that he did establish the sufficiency of certain orthogonality conditions.

$$m^{\frac{1}{2}}(L - \delta)/[Q/(m-1)]^{\frac{1}{2}} = t_{m-1},$$

where t_k denotes generically a variable having the t-distribution with k degrees of freedom. Define $t_{k,\epsilon}$ from

$$\Pr(-t_{k,\epsilon} \leqq t_k \leqq t_{k,\epsilon}) = \epsilon. \tag{1}$$

Then a set of confidence intervals for δ with confidence coefficient ϵ is

$$|\delta - L| \leqq t_{m-1,\epsilon}\{Q/[m(m-1)]\}^{\frac{1}{2}}. \tag{2}$$

Denote by $E(l)$ the expected length of the confidence interval (2),

$$E(l) = 2t_{m-1,\epsilon}[m(m-1)]^{-\frac{1}{2}}\sigma E[(Q/\sigma^2)^{\frac{1}{2}}],$$

$$E(l) = t_{m-1,\epsilon}c_{m-1}\sigma/m^{\frac{1}{2}}, \tag{3}$$

where

$$c_k = 2k^{-\frac{1}{2}}E(\chi_k) = (8/k)^{\frac{1}{2}}\Gamma(\tfrac{1}{2}k + \tfrac{1}{2})/\Gamma(\tfrac{1}{2}k).$$

The symmetrical choice (1) of the limits on the t-distribution minimizes (3).

We consider using in connection with the confidence intervals (2) linear functions

$$d_i = x_i - \sum_{j=1}^{n} c_{ij}y_j, \qquad i = 1, 2, \cdots, m. \tag{4}$$

The variables d_i have a multivariate normal distribution. Necessary and sufficient conditions that the d_i all have the same mean δ, equal variances σ^2, and zero covariances, are easily found to be

$$\sum_{j=1}^{n} c_{ij} = 1, \qquad \sum_{k=1}^{n} c_{ik}c_{jk} = c^2\delta_{ij}, \tag{5}$$

where $\delta_{ii} = 1$, $\delta_{ij} = 0$ if $i \neq j$. If (4) are used in (2), $E(l)$ is given by (3) with $\sigma^2 = \mu + c^2\nu$. Hence to minimize $E(l)$ we must find an $m \times n$ matrix $C = (c_{ij})$ satisfying (5), and for which c^2 is minimum. The minimum value of c^2 is m/n: this is easily proved by the use of vector algebra.

Let γ_i be the i-th row of C, and let ψ be the $1 \times n$ matrix $(1, 1, \cdots, 1)$. Denote the transpose of a matrix by a prime. Then the conditions (5) read

$$\gamma_i\psi' = 1, \qquad \gamma_i\gamma_j' = c^2\delta_{ij}. \tag{6}$$

First suppose vectors $\gamma_1, \gamma_2, \cdots, \gamma_m$ satisfy (6). Then it is possible to adjoin $n - m$ orthogonal vectors $\gamma_{m+1}, \cdots, \gamma_n$, so that the complete set satisfies the second group of conditions (6). Since this set is a basis in n-space,

$$\psi = \sum_{k=1}^{n} g_k\gamma_k,$$

where the g_k are scalars. Now

$$1 = \gamma_i \psi' = \sum_{k=1}^{n} g_k \gamma_i \gamma_k' = g_i c^2, \qquad i = 1, 2, \cdots, m,$$

and thus $g_1 = g_2 = \cdots = g_m = c^{-2}$. But

$$n = \psi\psi' = \sum_{k=1}^{n} g_k^2 \gamma_k \gamma_k' = mc^{-2} + \sum_{k=m+1}^{n} g_k^2 c^2 \geqq mc^{-2},$$

and hence $c^2 \geqq m/n$. On the other hand this lower bound for c^2 may be attained by taking any set $\beta_1, \beta_2, \cdots, \beta_m$ of orthogonal vectors with norms m/n, that is, $\beta_i \beta_j' = \delta_{ij} m/n$, and rotating them so that their equal angles vector $\lambda = (n/m)(\beta_1 + \beta_2 + \cdots + \beta_m)$ coincides with ψ. Then $\lambda S = \psi$, where $SS' = I$. For $\gamma_i = \beta_i S$,

$$\gamma_i \psi' = \beta_i SS'\lambda' = \beta_i \lambda' = 1,$$

$$\gamma_i \gamma_j' = \beta_i SS' \beta_j' = \beta_i \beta_j' = \delta_{ij} m/n,$$

so that equations (6) are satisfied with $c^2 = m/n$.

An especially neat solution of this minimum problem was obtained by the above method; its validity may easily be verified directly. It is

$$c_{ij} = \begin{cases} \delta_{ij}(m/n)^{\frac{1}{2}} - (mn)^{-\frac{1}{2}} + 1/n, & j \leqq m, \\ 1/n, & j > m. \end{cases}$$

Then

$$d_i = x_i - (m/n)^{\frac{1}{2}} y_i + (mn)^{-\frac{1}{2}} \sum_{j=1}^{m} y_j - \sum_{j=1}^{n} y_j/n,$$

and L and Q become simply $L = \bar{x} - \bar{y}$ and

$$Q = \sum_{i=1}^{m} (u_i - \bar{u})^2, \tag{7}$$

where

$$\bar{x} = \sum_{i=1}^{m} x_i/m, \qquad \bar{y} = \sum_{i=1}^{n} y_i/n, \qquad u_i = x_i - (m/n)^{\frac{1}{2}} y_i, \qquad \bar{u} = \sum_{i=1}^{m} u_i/m. \tag{8}$$

We may now write (2) as[3]

$$\bar{x} - \bar{y} - t_{m-1,\epsilon}\{Q/[m(m-1)]\}^{\frac{1}{2}} \leqq \delta \leqq \bar{x} - \bar{y} + t_{m-1,\epsilon}\{Q/[m(m-1)]\}^{\frac{1}{2}}. \tag{9}$$

The solution (B) mentioned at the beginning, consists of taking $c_{ij} = \delta_{ij}$ in (4), so that the conditions (5) are satisfied with $c^2 = 1$. Hence for both (B) and (9) the expected length of the confidence interval is given by (3), but with $\sigma^2 = \mu + \nu$ for (B), while $\sigma^2 = \mu + (m/n)\nu$ for (9).

[3] Obvious modifications of (9) will make it suitable for "one-sided" estimation.

3. More General Solutions. We now generalize our approach to the following extent: Let L be a linear form and Q a quadratic form in the variates $x_1, \cdots, x_m, y_1, \cdots, y_n$, with coefficients independent of the parameters (i. of p.). If for some constant h i. of p., and some function f of the parameters, $h(L - \delta)/f$ and Q/f^2 are independently distributed, the former according to the normal law with zero mean and unit variance, the latter according to the χ^2-law with $k - 1$ degrees of freedom, then the quotient

$$h(L - \delta)/[Q/(k - 1)]^{\frac{1}{2}} \tag{10}$$

will have the t-distribution with $k - 1$ degrees of freedom, no matter what the values of the parameters.

We note that necessarily then

$$E(L) = \delta, \tag{11}$$

$$f^2 = h^2 E[(L - \delta)^2]. \tag{12}$$

The t-distribution of (10) leads to the confidence intervals

$$| \delta - L | \leqq t_{k-1,\epsilon}[Q/(k - 1)]^{\frac{1}{2}}/h, \tag{13}$$

where $t_{k-1,\epsilon}$ is defined by (1), and the confidence coefficient is ϵ. Proceeding as toward (3), we find that the expected length of (13) is

$$E(l) = t_{k-1,\epsilon} c_{k-1} f/h. \tag{14}$$

If

$$L = \sum_{i=1}^{m} a_i x_i - \sum_{i=1}^{n} b_i y_i,$$

$$E(L) = \alpha \sum_{i=1}^{m} a_i - \beta \sum_{i=1}^{n} b_i. \tag{15}$$

Since a_i, b_i are i. of p., it follows from (11) and (15) that

$$\sum_{i=1}^{m} a_i = \sum_{i=1}^{n} b_i = 1. \tag{16}$$

Writing

$$\xi_i = x_i - \alpha, \qquad \eta_i = y_i - \beta,$$

$$L - \delta = \sum_{i=1}^{m} a_i \xi_i - \sum_{i=1}^{n} b_i \eta_i, \tag{17}$$

$$E[(L - \delta)^2] = \mu \sum_{i=1}^{m} a_i^2 + \nu \sum_{i=1}^{n} b_i^2 = f^2/h^2 \tag{18}$$

from (12); thus (14) may be written

$$E(l) = t_{k-1,\epsilon} c_{k-1} \left[\mu \sum_{i=1}^{m} a_i^2 + \nu \sum_{i=1}^{n} b_i^2 \right]^{\frac{1}{2}}. \tag{19}$$

From (18) we also have

$$f^2 = a^2 \mu + b^2 \nu, \tag{20}$$

where

$$a^2 = h^2 \sum_{i=1}^{m} a_i^2, \qquad b^2 = h^2 \sum_{i=1}^{n} b_i^2$$

are i. of p.

4. Lemma. We propose to prove next that the maximum value of k is m, that is to say, it is impossible to obtain a t-distribution for a quotient (10) with more than $m - 1$ degrees of freedom. For this we need a lemma to the effect that certain well known sufficient conditions for a quadratic form to have a χ^2-distribution are also necessary.

Since under our assumptions $h^2(L - \delta)^2/f^2 = \chi_1^2$ and $Q/f^2 = \chi_{k-1}^2$ are independent, therefore $Q^*/f^2 = \chi_k^2$, where

$$Q^* = h^2(L - \delta)^2 + Q.$$

To shorten the notation, write

$$z_i = \begin{cases} x_i, & i = 1, 2, \cdots, m, \\ y_{i-m}, & i = m + 1, \cdots, m + n, \end{cases}$$

$$\alpha_i = E(z_i), \qquad \zeta_i = z_i - \alpha_i, \qquad \sigma_i^2 = E(\zeta_i^2).$$

Let

$$Q = \sum_{s,t} q_{st} z_s z_t,$$

where the indices s and t range from 1 to $m + n$ throughout. Then q_{st} is i. of p., and

$$Q = \sum_{s,t} q_{st} \zeta_s \zeta_t + 2 \sum_{s} q_s \zeta_s + q,$$

where

$$q_s = \sum_{t} q_{st} \alpha_t, \qquad q = \sum_{s} q_s \alpha_s.$$

From (17)

$$h^2(L - \delta)^2 = \sum_{s,t} p_{st} \zeta_s \zeta_t,$$

where p_{st} are i. of p. Putting $q_{st}^* = q_{st} + p_{st}$, q_{st}^* are i. of p., and

$$Q^* = \sum_{s,t} q_{st}^* \zeta_s \zeta_t + 2 \sum_{s} q_s \zeta_s + q. \tag{21}$$

The moment-generating function of Q^*/f^2 is

$$\phi(\theta) = E[\exp(\theta Q^*/f^2)] = C_1 \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \exp\left(\theta Q^*/f^2 - \tfrac{1}{2} \sum_{s} \zeta_s^2/\sigma_s^2\right) \prod_{s} d\zeta_s.$$

There exists a non-singular linear transformation from the ζ's to v's such that

$$\sum_s \zeta_s^2/\sigma_s^2 = \sum_s v_s^2 ,$$

$$\sum_{s,t} q_{st}^* \zeta_s \zeta_t = \sum_s \lambda_s v_s^2 . \tag{22}$$

Then

$$\sum_s q_s \zeta_s = \sum_s p_s v_s , \tag{23}$$

$$\begin{aligned}\phi(\theta) &= C_2 e^{\theta q/f^2} \prod_s \int_{-\infty}^{+\infty} \exp\{-\tfrac{1}{2}[v^2 - 2\theta(\lambda_s v^2 + 2p_s v)/f^2]\}\, dv \\ &= e^{\theta q/f^2} \prod_s (1 - 2\theta\lambda_s/f^2)^{-\frac{1}{2}} \exp\{2\theta^2 p_s^2/(f^4 - 2\theta\lambda_s f^2)\}.\end{aligned}$$

Now $Q^*/f^2 = \chi_k^2$ if and only if

$$\phi(\theta) = (1 - 2\theta)^{-\frac{1}{2}k}.$$

Hence

$$p_s = 0, \qquad q = 0, \tag{24}$$

and k of the λ_s must be equal to f^2 while the remaining λ_s vanish. No generality is lost in assuming

$$\lambda_1 = \lambda_2 = \cdots = \lambda_k = f^2, \qquad \lambda_{k+1} = \cdots = \lambda_{m+n} = 0. \tag{25}$$

Let $w_i = fv_i$, $i = 1, 2, \cdots, k$. From equations (21) to (25) we deduce that

$$Q^* = \sum_{s,t} q_{st}^* \zeta_s \zeta_t = \sum_{i=1}^{k} w_i^2 , \tag{26}$$

where q_{st}^* is i. of p., and the w_i are linear combinations of the ζ_s such that

$$E(w_i w_j) = f^2 \delta_{ij} . \tag{27}$$

That the conditions (26) are necessary[4] for $Q^*/f^2 = \chi_k^2$ constitutes the desired lemma.

5. Maximum Number of Degrees of Freedom. We have seen that the w_i in (26) must be of the form

$$w_i = \sum_{j=1}^{m} a_{ij} \xi_j + \sum_{j=1}^{n} b_{ij} \eta_j . \tag{28}$$

We substitute (28) and (20) into (27) and write the result in matrix form,

$$\mu AA' + \nu BB' = (a^2\mu + b^2\nu) I_k , \tag{29}$$

where I_j is the identity matrix of order j, $A^{k\times m} = (a_{ij})$, $B^{k\times n} = (b_{ij})$, and whenever a new matrix is introduced, a superscript $r \times c$ indicates that it has r rows

[4] We have incidentally proved sufficiency.

and c columns. Now if we knew that AA' and BB' were i. of p., then we could equate coefficients of μ and ν in (29) and immediately draw the desired conclusion $k \leqq m$. But that AA', BB' are i. of p. is not obvious, since this need not be true of A and B. However, we do know that the matrices

$$F = A'A, \qquad G = A'B, \qquad H = B'B$$

are i. of p. because the matrix (q^*_{st}) of (26) is

$$\begin{pmatrix} F & G \\ G' & H \end{pmatrix}.$$

Multiplying (29) on the left by A' and on the right by A, we obtain

$$\mu F^2 + \nu GG' = (a^2\mu + b^2\nu)F. \tag{30}$$

(30) must hold identically in μ, ν. Since the coefficients of μ, ν are now i. of p., we may equate them, hence $GG' = b^2F$. Similarly multiplying (29) by B' and B, we get $G'G = a^2H$. Now[5] for any matrix M, rank M = rank $M'M$ = rank MM'. Thus rank F = rank $H = r$, say. Again, $F = A'A$, therefore r = rank $A \leqq m$. Since F is a positive[5] matrix, i. of p., there exists a non-singular $P^{m\times m}$, i. of p., such that

$$F = P'I_{m,r}P = A'A, \tag{31}$$

where $I_{j,r}$ is the $j \times j$ matrix the first r of whose diagonal elements are unity and all other elements zero. Let $T^{k\times m} = AP^{-1}$. Then

$$A = TP, \qquad T'T = I_{m,r} \tag{32}$$

from (31). Likewise we can write

$$B = U^{k\times n}R^{n\times n}, \qquad U'U = I_{n,r}, \tag{33}$$

where R is non-singular and i. of p. Then $G = A'B = P'T'UR$, hence $T'U = (P')^{-1}GR^{-1}$ is i. of p. We note

$$T = (T_1^{k\times r}, 0^{k\times(m-r)}), \qquad U = (U_1^{k\times r}, 0^{k\times(n-r)}),$$

where

$$T_1'T_1 = U_1'U_1 = I_r.$$

Since

$$T'U = \begin{pmatrix} T_1' \\ 0 \end{pmatrix}(U_1, 0) = \begin{pmatrix} T_1'U_1 & 0 \\ 0 & 0 \end{pmatrix}$$

---

[5] A simple proof [5b] of this useful theorem is the following: Let r = rank M, p = rank $M'M$. $p \leqq r$ since the rank of the product cannot exceed the rank of a factor. M contains r independent column vectors; the Grammian matrix of these vectors is non-singular and appears as an $r \times r$ minor in $M'M$. Hence $p \geqq r$. Furthermore, all principal minors of $M'M$ are Grammian matrices (which always have non-negative determinants), hence $M'M$ is always positive——we use this below.

is i. of p., so is its minor $V^{r\times r} = T_1'U_1$. Write

$$P = \begin{pmatrix} P_1^{r\times m} \\ P_2^{(m-r)\times m} \end{pmatrix}, \qquad R = \begin{pmatrix} R_1^{r\times n} \\ R_2^{(n-r)\times n} \end{pmatrix}.$$

Then from (32), (33),

$$A = T_1P_1, \qquad B = U_1R_1. \tag{34}$$

Substituting (34) in (29), we get

$$\mu T_1P_1P_1'T_1' + \nu U_1R_1R_1'U_1' = (a^2\mu + b^2\nu)I_k, \tag{35}$$

and multiplying by T_1' on the left, T_1 on the right,

$$\mu P_1P_1' + \nu VR_1R_1'V' = (a^2\mu + b^2\nu)I_r.$$

Again the coefficients of μ, ν are i. of p., so

$$\begin{aligned} P_1P_1' &= a^2I_r, \\ VR_1R_1'V' &= b^2I_r. \end{aligned} \tag{36}$$

Similarly we find

$$R_1R_1' = b^2I_r. \tag{37}$$

From (36), (37), $VV' = I_r$. (35) now becomes

$$a^2\mu T_1T_1' + b^2\nu U_1U_1' = (a^2\mu + b^2\nu)I_k. \tag{38}$$

Multiplication of (38) on the right by U_1 gives

$$a^2\mu T_1V + b^2\nu U_1 = (a^2\mu + b^2\nu)U_1.$$

Hence $T_1V = U_1$, therefore $U_1U_1' = T_1T_1'$, and putting this back into (38) we have $I_k = T_1T_1'$, rank I_k = rank T_1T_1' = rank $T_1'T_1$ = rank I_r, $k = r \leqq m$.

6. Minimum Expected Length of Confidence Intervals. We now point out that of all confidence intervals (13) with $k = m$, the confidence intervals (9) have the minimum expected length. Recalling that the a_i, b_i in (19) are subject to the conditions (16), we easily find

$$\sum_{i=1}^{m} a_i^2 \geqq 1/m, \qquad \sum_{i=1}^{n} b_i^2 \geqq 1/n. \tag{39}$$

From (39) and (19) we have

$$E(l) \geqq t_{m-1,\epsilon}c_{m-1}[\mu + (m/n)\nu]^{\frac{1}{2}}/m^{\frac{1}{2}},$$

and referring to the statement at the end of section 2, the property of (9) asserted above is now obvious.

7. Asymptotic Shortness of Confidence Intervals. In conclusion we wish to compare our results with the case where the ratio of the variances, $\theta = \nu/\mu$, is known. If

$$S_x = \sum_{i=1}^{m} (x_i - \bar{x})^2, \qquad S_y = \sum_{i=1}^{n} (y_i - \bar{y})^2,$$

$$L = \bar{x} - \bar{y}, \qquad \sigma_L^2 = (\mu/m) + (\nu/n),$$

then $(L - \delta)/\sigma_L$, S_x/μ, S_y/ν are mutually independently distributed, the first normally with zero mean and unit variance, and $S_x/\mu = \chi^2_{m-1}$, $S_y/\nu = \chi^2_{n-1}$. Hence

$$(L - \delta)[(\mu m^{-1} + \nu n^{-1})(S_x\mu^{-1} + S_y\nu^{-1})/(m + n - 2)]^{-\frac{1}{2}} = t_{m+n-2}.$$

TABLE I

Values of R for $\epsilon = .95$

$m-1$ \ $n-1$	5	10	20	40	∞
5	1.15	1.20	1.23	1.25	1.28
10		1.05	1.07	1.09	1.11
20			1.03	1.03	1.05
40				1.01	1.02

TABLE II

Values of R for $\epsilon = .99$

$m-1$ \ $n-1$	5	10	20	40	∞
5	1.27	1.36	1.42	1.47	1.52
10		1.10	1.13	1.16	1.20
20			1.05	1.06	1.09
40				1.02	1.04

This relation yields the confidence intervals

$$(40) \qquad |\delta - L| \leqq t_{m+n-2,\epsilon}(m + n - 2)^{-\frac{1}{2}}(m^{-1} + \theta n^{-1})^{\frac{1}{2}}(S_x + S_y/\theta)^{\frac{1}{2}},$$

where the confidence coefficient is again ϵ. The confidence intervals (40) are known to be highly efficient; for instance they are of the shortest unbiased type [5a]. We calculate their expected length to be

$$E(l) = t_{m+n-2,\epsilon}c_{m+n-2}[\mu + (m/n)\nu]^{\frac{1}{2}}/m^{\frac{1}{2}}.$$

The ratio R of $E(l)$ for (9) to $E(l)$ for (40) is thus

$$(41) \qquad R = (t_{m-1,\epsilon}c_{m-1})/(t_{m+n-2,\epsilon}c_{m+n-2}).$$

As $k \to \infty$, $c_k \to 2$, $t_{k,\epsilon} \to t_{\infty,\epsilon}$, hence as $m \to \infty$, $R \to 1$ no matter what the values of $n \geqq m$. For small values of m the ratio of the t values in (41) is considerably >1, but this is partly offset by c_k approaching its limiting value 2 from below so that the ratio of the c's is <1. The behaviour of R for finite m is indicated in Tables I and II. Table I (II) tells us for example that with $m > 10$, and $\epsilon = .95$ (.99), the expected length of the confidence intervals (9) is at most 11 per cent (20%) longer than that of the optimum confidence intervals (40) available when the ratio θ is known. While we may conclude from $R \to 1$ as $m \to \infty$, that our solution (9) is asymptotically extremely efficient, we cannot conclude from Tables I, II that for small m (9) is inefficient, since we do not know what the lengthening effect of the extra nuisance parameter in the Behrens-Fisher problem would be on "best" confidence intervals.

8. Summary. In the terminology of the first paragraphs of sections 1 and 3 we have proved that there do not exist a linear form L and a quadratic form Q in the observations such that the quotient (10) will have the t-distribution (for *all* values of the parameters) with more than $m - 1$ degrees of freedom. We have further shown that of all confidence intervals (13) based on the t-distribution with $m - 1$ degrees of freedom, and with confidence coefficient ϵ, (9) has the minimum expected length. The quantities needed to apply our solution (9) are given by (1), (7) and (8). Finally, by comparing this solution with a known highly efficient solution for the case when the ratio of the population variances is known, it has been possible to show that at least asymptotically our confidence intervals (9) are very short.

REFERENCES

[1] W. U. Behrens, *Landwirtschaftliche Jahrbücher*, Berlin, Vol. 68 (1929), pp. 807–837.
[2] R. A. Fisher, *Annals of Eugenics*, Vol. 6 (1935), pp. 391–398; Vol. 9 (1939), pp. 174–180; Vol. 11 (1941), pp. 141–172. *Annals of Math. Stat.*, Vol. 10 (1939), pp. 383–388.
[3] J. Neyman, *Biometrika*, Vol. 32 (1941), p. 138.
[4] B. L. Welch, *Biometrika*, Vol. 29 (1938), p. 360.
[5] H. Scheffé, [a] *Annals of Math. Stat.*, Vol. 13 (1942), p. 386. [b] *Amer. Math. Monthly*, Vol. 49 (1942), p. 101.

Further Bibliography

[6] M. S. Bartlett, *Proc. Camb. Phil. Soc.*, Vol. 32 (1936), pp. 560–566. *Annals of Math. Stat.*, Vol. 10 (1939) pp. 129–138.
[7] P. L. Hsu, *Stat. Res. Mem.*, Vol. 2 (1938), pp. 1–25.
[8] H. Jeffreys, *Annals of Eugenics*, Vol. 10 (1940), pp. 48–51.
[9] E. J. G. Pitman, *Suppl. J. Roy. Stat. Soc.*, Vol. 4 (1937), pp. 117–130.
[10] D. M. Starkey, *Annals of Math. Stat.*, Vol. 9 (1938), pp. 201–213.
[11] P. V. Sukhatme, *Sankhyā*, Vol. 4 (1938), pp. 39–48.
[12] S. S. Wilks, *Annals of Math. Stat.*, Vol. 11 (1940), p. 474.
[13] F. Yates, *Proc. Camb. Phil. Soc.*, Vol. 35 (1939), pp. 579–609.

AN EXTENSION OF WILKS' METHOD FOR SETTING TOLERANCE LIMITS

By Abraham Wald

Columbia University

1. Introduction. Let x be a random variable and let $f(x)$ be its probability density function. Suppose that nothing is known about $f(x)$ except that it is continuous. Let $x_1, \cdots, x_n$ be n independent observations on x. The problem of setting tolerance limits can be formulated as follows: *For some given positive values $\beta < 1$ and $\gamma < 1$ we have to construct two functions $L(x_1, \cdots, x_n)$ and $M(x_1, \cdots, x_n)$, called tolerance limits, such that the probability that*

$$\int_L^M f(t)\,dt \geq \gamma, \tag{1}$$

holds, is equal to β. This problem has recently been solved by S. S. Wilks[1] in a very satisfactory way when nothing is known about $f(x)$ except that it is continuous. Wilks proposes the following solution: Let $x_1, \cdots, x_n$ be the observed values of x arranged in order of increasing magnitude. Then $L = x_r$ and $M = x_{n-r+1}$ where r denotes a positive integer. The exact sampling distribution of the statistic $\int_{x_r}^{x_{n-r+1}} f(t)\,dt$ is derived by Wilks and this provides the solution for the problem of setting tolerance limits. A very important feature of Wilks' solution is the fact that the distribution of $\int_{x_r}^{x_{n-r+1}} f(t)\,dt$ is entirely independent of the unknown density function $f(x)$, i.e. the distribution of $\int_{x_r}^{x_{n-r+1}} f(t)\,dt$ is the same for any arbitrary continuous density function $f(x)$.

In this paper we shall give an extension of Wilks' method to the multivariate case. Let $x_1, \cdots, x_p$ be a set of p random variables with the joint probability density function $f(x_1, \cdots, x_p)$. Suppose that nothing is known about $f(x_1, \cdots, x_p)$ except that it is a continuous function of $x_1, \cdots, x_p$. A sample of n independent observations is drawn and the α-th observation on x_i is denoted by $x_{i\alpha}$ $(i = 1, \cdots, p;\ \alpha = 1, \cdots, n)$. The problem of setting tolerance limits for $x_1, \cdots, x_p$ can be formulated as follows: *For some given positive values $\beta < 1$ and $\gamma < 1$ we have to construct p pairs of functions of the observations $L_i(x_{11}, \cdots, x_{pn})$ and $M_i(x_{11}, \cdots, x_{pn})$ $(i = 1, \cdots, p)$ such that the probability that*

$$\int_{L_p}^{M_p} \cdots \int_{L_1}^{M_1} f(t_1, \cdots, t_p)\,dt_1 \cdots dt_p \geq \gamma, \tag{2}$$

[1] S. S. Wilks, "Determination of sample sizes for setting tolerance limits," *Annals of Math. Stat.*, Vol. 12 (1941).

holds, is equal to β. The functions L_i and M_i are called the lower and upper tolerance limits of x_i. A natural extension of Wilks' procedure would seem to be the following: Let $x_{i1}, \cdots, x_{in}$ be the observations on x_i arranged in order of increasing magnitude and let $L_i = x_{ir_i}$ and $M_i = x_{is_i}$ $(i = 1, \cdots, p)$ where r_i and s_i denote some integers. However, this choice of the tolerance limits does not provide a satisfactory solution of our problem, since the distribution of (2) is *not* independent of the unknown density function $f(x_1, \cdots, x_p)$. It will be shown in this paper that by a slight modification of the above procedure the distribution of (2) becomes entirely independent of the unknown density function $f(x_1, \cdots, x_p)$. In section 2 we will treat the bivariate case and in section 3 we will extend the results to multivariate distributions.

2. The bivariate case. In this section we deal with the case when $p = 2$. Let $x_{11}, \cdots, x_{1n}$ be the observations on x_1 arranged in order of increasing magnitude. We may assume that $x_{11} < x_{12} < \cdots < x_{1n}$ since the probability of an equality sign is equal to zero. We define

$$L_1 = x_{1r_1} \quad \text{and} \quad M_1 = x_{1s_1}, \tag{3}$$

where r_1 and s_1 denote some positive integers and $r_1 < s_1 \leq n$. Consider only those sample points $(x_{1\alpha}, x_{2\alpha})$ for which $x_{1r_1} < x_{1\alpha} < x_{1s_1}$, i.e. consider the sample points $(x_{1,r_1+1}, x_{2,r_1+1}), \cdots, (x_{1,s_1-1}, x_{2,s_1-1})$. Denote by $x'_{2,r_1+1}, \cdots, x'_{2,s_1-1}$ the values $x_{2,r_1+1}, \cdots, x_{2,s_1-1}$ arranged in order of increasing magnitude. We define

$$L_2 = x'_{2r_2} \quad \text{and} \quad M_2 = x'_{2s_2}, \tag{4}$$

where r_2 and s_2 denote some positive integers for which $r_2 < s_2 \leq s_1 - r_1 - 1$.

We will show that the distribution of the statistic

$$Q = \int_{L_2}^{M_2} \int_{L_1}^{M_1} f(t_1, t_2)\, dt_1\, dt_2, \tag{5}$$

is entirely independent of the unknown density function $f(x_1, x_2)$. Denote by $\varphi(x_1)$ the marginal distribution of x_1, i.e.

$$\varphi(x_1) = \int_{-\infty}^{+\infty} f(x_1, x_2)\, dx_2. \tag{6}$$

Furthermore denote by $\psi(x_2, L_1, M_1)$ the conditional distribution of x_2 calculated under the condition that $L_1 < x_1 < M_1$. Hence

$$\psi(x_2, L_1, M_1) = \frac{\displaystyle\int_{L_1}^{M_1} f(x_1, x_2)\, dx_1}{\displaystyle\int_{-\infty}^{+\infty} \int_{L_1}^{M_1} f(x_1, x_2)\, dx_1\, dx_2}. \tag{7}$$

Let

$$P = \int_{L_1}^{M_1} \varphi(t)\, dt \tag{8}$$

and

$$\overline{P} = \int_{L_2}^{M_2} \psi(t, L_1, M_1)\, dt \tag{9}$$

From (5), (8) and (9) it follows that

$$Q = P\overline{P}. \tag{10}$$

It is obvious that the distribution of P is given by Wilks' formula. Since Wilks derived the distribution only when $s_1 = n - r_1 + 1$, we will briefly give here the derivation for any integers r_1 and s_1.

Let $\int_{-\infty}^{x_{1r_1}} \varphi(t)\, dt = u$ and $\int_{x_{1s_1}}^{\infty} \varphi(t)\, dt = v$. Then the joint probability density function of u and v is given by

$$cu^{r_1-1}(1 - u - v)^{s_1-r_1-1} v^{n-s_1}\, du\, dv, \tag{11}$$

where c is a constant. We obviously have $P = 1 - u - v$. The joint density function of P and u is given by

$$cu^{r_1-1}P^{s_1-r_1-1}(1 - u - P)^{n-s_1}\, du\, dP, \tag{12}$$

where u is restricted to the interval $[0, 1 - P]$. Hence the distribution of P is given by

$$\begin{aligned} &cP^{s_1-r_1-1} \int_0^{1-P} u^{r_1-1}(1 - u - P)^{n-s_1}\, du \\ &\quad = cP^{s_1-r_1-1}(1 - P)^{n-s_1+r_1-1} \int_0^{1-P} \left(\frac{u}{1-P}\right)^{r_1-1} \left(1 - \frac{u}{1-P}\right)^{n-s_1} du \\ &\quad = cP^{s_1-r_1-1}(1 - P)^{n-s_1+r_1} \int_0^1 T^{r_1-1}(1 - T)^{n-s_1}\, dT \\ &\quad = c'P^{s_1-r_1-1}(1 - P)^{n-s_1+r_1}. \end{aligned}$$

Since the integral of the density function of P over the range of P must be equal to 1, we find that

$$c' = \Gamma(n + 1)/\Gamma(s_1 - r_1)\Gamma(n - s_1 + r_1 + 1).$$

Hence the probability density function of P is given by

$$\frac{\Gamma(n + 1)}{\Gamma(s_1 - r_1)\Gamma(n - s_1 + r_1 + 1)} P^{s_1-r_1-1}(1 - P)^{n-s_1+r_1}\, dP. \tag{13}$$

Since $x_{2,r_1+1}, \cdots, x_{2,s_1-1}$ can be considered as $s_1 - r_1 - 1$ independent observations on a random variable t the distribution of which is given by $\psi(t, L_1, M_1)\, dt$, for any given values L_1, M_1 the conditional distribution of $\overline{P}$ is given by the expression we obtain from (13) by substituting r_2 for r_1, s_2 for s_1 and $s_1 - r_1 - 1$ for n. Hence the conditional distribution of $\overline{P}$ is given by

$$(14) \qquad \frac{\Gamma(s_1 - r_1)}{\Gamma(s_2 - r_2)\Gamma(s_1 - r_1 - s_2 + r_2)} (\overline{P})^{s_2 - r_2 - 1}(1 - \overline{P})^{s_1 - r_1 - 1 - s_2 + r_2} \, d\overline{P}.$$

Since the expression (14) does not involve the quantities L_1 and M_1, $\overline{P}$ is distributed independently of L_1 and M_1. Hence the joint density function of P and $\overline{P}$ is given by the product of (13) and (14), i.e. by

$$(15) \qquad AP^{s_1 - r_1 - 1}(1 - P)^{n - s_1 + r_1}(\overline{P})^{s_2 - r_2 - 1}(1 - \overline{P})^{s_1 - r_1 - 1 - s_2 + r_2} \, dP \, d\overline{P},$$

where A denotes the product of the constant coefficients in (13) and (14). From (15) it follows that the joint distribution of P and $Q = P\overline{P}$ is given by

$$(16) \qquad A(1 - P)^{n - s_1 + r_1} Q^{s_2 - r_2 - 1}(P - Q)^{s_1 - r_1 - 1 - s_2 + r_2} \, dP \, dQ.$$

Since the range of P is the interval $[Q, 1]$, the distribution of Q is given by

$$(17) \qquad AQ^{s_2 - r_2 - 1} \int_Q^1 (1 - P)^{n - s_1 + r_1}(P - Q)^{s_1 - r_1 - 1 - s_2 + r_2} \, dP.$$

Let $R = P - Q$. Then we have

$$\begin{aligned} &\int_Q^1 (1 - P)^{n - s_1 + r_1}(P - Q)^{s_1 - r_1 - 1 - s_2 + r_2} \, dP \\ (18) \qquad &= \int_0^{1 - Q} (1 - Q - R)^{n - s_1 + r_1} R^{s_1 - r_1 - 1 - s_2 + r_2} \, dR \\ &= (1 - Q)^{n - 1 - s_2 + r_2}(1 - Q) \int_0^1 (1 - T)^{n - s_1 + r_1} T^{s_1 - r_1 - 1 - s_2 + r_2} \, dT. \end{aligned}$$

From (17) and (18) it follows that the probability density function of Q is given by

$$(19) \qquad \frac{\Gamma(n + 1)}{\Gamma(s_2 - r_2)\Gamma(n - s_2 + r_2 + 1)} Q^{s_2 - r_2 - 1}(1 - Q)^{n - s_2 + r_2} \, dQ.$$

3. The multivariate case. We may assume that no two elements of the matrix $\| x_{i\alpha} \|$ $(i = 1, \cdots, p; \alpha = 1, \cdots, n)$ are equal, since the probability of this event is equal to 1. For each α let $t_\alpha (\alpha = 1, \cdots, n)$ be the point with the coordinates $x_{1\alpha}, \cdots, x_{p\alpha}$. Let $x_{11}, \cdots, x_{1n}$ be the observations on x_1 arranged in order of increasing magnitude. Then $L_1 = x_{1r_1}$ and $M_1 = x_{1s_1}$. The quantities L_i and M_i $(i = 2, \cdots, p)$ are defined in the following manner: Let S be the set of all points t_α for which

$$L_j < x_{j\alpha} < M_j \qquad (j = 1, \cdots, i - 1).$$

Arrange the i-th coordinates of the points in S in order of increasing magnitude. Then L_i is equal to the r_i-th element and M_i is equal to the s_i-th element of this ordered sequence. We will derive the distribution of

$$(20) \qquad Q_p = \int_{L_p}^{M_p} \cdots \int_{L_1}^{M_1} f(x_1, \cdots, x_p) \, dx_1 \cdots dx_p.$$

Let

$$Q_i = \int_{-\infty}^{+\infty} \cdots \int_{-\infty}^{+\infty} \int_{L_i}^{M_i} \cdots \int_{L_1}^{M_1} f(x_1, \cdots, x_p)\, dx_1 \cdots dx_p \tag{21}$$

$$(i = 1, \cdots, p-1).$$

Denote by $\varphi_i(x_i, L_1, M_1, \cdots, L_{i-1}, M_{i-1})$ $(i = 2, \cdots, p)$ the conditional probability density function of x_i calculated under the condition that $L_j \leq x_j \leq M_j$ $(j = 1, \cdots, i-1)$. Let

$$\overline{P}_i = \int_{L_i}^{M_i} \varphi_i(x_i, L_1, M_1, \cdots, L_{i-1}, M_{i-1})\, dx_i. \tag{22}$$

We obviously have

$$Q_{i+1} = Q_i \overline{P}_{i+1} \qquad (i = 1, \cdots, p-1). \tag{23}$$

We will prove that the probability density function of Q_i is given by

$$\frac{\Gamma(n+1)}{\Gamma(s_i - r_i)\Gamma(n - s_i + r_i + 1)} Q_i^{s_i - r_i - 1}(1 - Q_i)^{n - s_i + r_i}\, dQ_i, \quad (i = 1, \cdots, p). \tag{24}$$

This is certainly true for $i = 1, 2$. We will assume that it is true for $i = j$ and we will prove it for $i = j + 1$. It is easy to see that Q_j and $\overline{P}_{j+1}$ are independently distributed and that the probability density function of $\overline{P}_{j+1}$ is given by

$$\frac{\Gamma(s_j - r_j)}{\Gamma(s_{j+1} - r_{j+1})\Gamma(s_j - r_j - s_{j+1} + r_{j+1})} \cdot (\overline{P}_{j+1})^{s_{j+1} - r_{j+1} - 1}(1 - \overline{P}_{j+1})^{s_j - r_j - 1 - s_{j+1} + r_{j+1}}\, d\overline{P}_{j+1}. \tag{25}$$

The joint distribution of Q_j and $\overline{P}_{j+1}$ is of the same form as the joint distribution of P and $\overline{P}$ in section 2. Hence the distribution of $Q_j \overline{P}_{j+1}$ can be obtained from the distribution of $Q = P\overline{P}$ by substituting r_{j+1} for r_2 and s_{j+1} for s_2. The distribution of Q is given in (19). Making the above substitution in formula (19) we obtain formula (24) for $i = j + 1$. Hence the validity of (24) is proved for $i = 1, 2, \cdots, p$. In particular, the distribution of Q_p is given by

$$\frac{\Gamma(n+1)}{\Gamma(s_p - r_p)\Gamma(n - s_p + r_p + 1)} Q_p^{s_p - r_p - 1}(1 - Q_p)^{n - s_p + r_p}\, dQ_p. \tag{26}$$

It is interesting to note that the distribution of Q_p does not depend on the integers $r_1, s_1, \cdots, r_{p-1}, s_{p-1}$. The construction of the tolerance limits L_i, M_i $(i = 1, \cdots, p)$, as proposed here, is somewhat asymmetric, since it depends on the order of the variates $x_1, \cdots, x_p$. In practical applications the asymmetry of the construction will be very slight, since in most practical cases the integers r_p and s_p will be chosen so that $(s_p - r_p - 1)/n$ will be near to 1. If, for example, $(s_p - r_p - 1)/n \geq .95$, the tolerance limits will be affected only very slightly by a permutation of the variates $x_1, \cdots, x_p$. However, it would be desirable to find a construction which is entirely independent of the order of the variates $x_1, \cdots, x_p$.

4. Tolerance regions composed of several rectangles. For the sake of simplicity we will consider here the bivariate case. All results obtained in this section can be extended without any difficulty to the multivariate case.

In section 2 the tolerance region has been a single rectangle in the plane (x_1, x_2) determined by the four lines $x_1 = L_1$, $x_1 = M_1$; $x_2 = L_2$ and $x_2 = M_2$. If the variates x_1 and x_2 are strongly correlated, a tolerance region of rectangle shape seems to be unfavorable, since it will cover an unnecessarily large area in the (x_1, x_2) plane. The situation is illustrated in figure 1 where the scatter of a

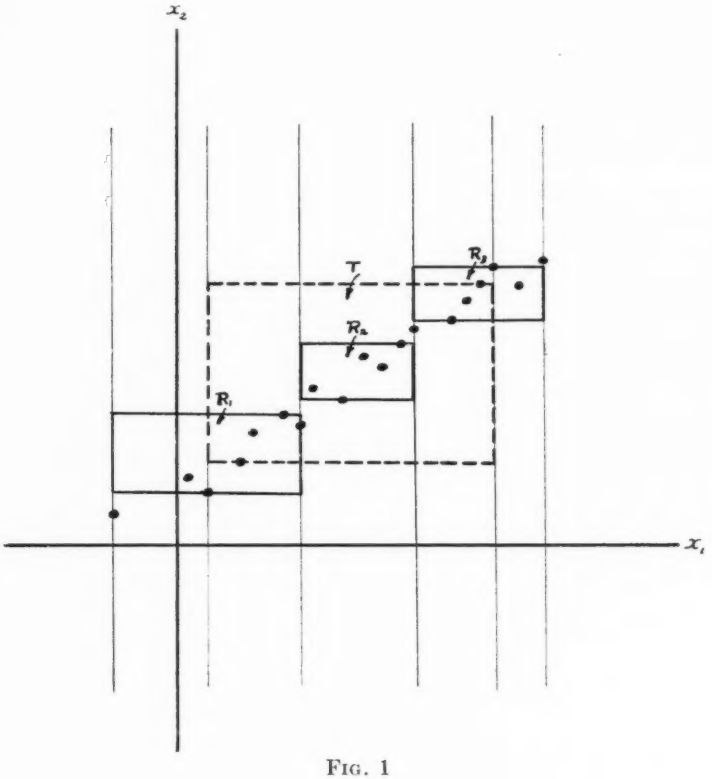

FIG. 1

bivariate sample of size $n = 19$ is shown. Suppose we choose $r_1 = 3$, $s_1 = 17$; $r_2 = 1$, $s_2 = 13$, then the tolerance region T, as defined in section 2, will be the rectangle determined by the lines $x_1 = L_1 = x_{1,3}$; $x_1 = M_1 = x_{1,17}$; $x_2 = L_2 = x'_{2,1}$; and $x_2 = M_2 = x'_{2,13}$. Now consider the tolerance region T' consisting of 3 small rectangles R_1, R_2 and R_3 defined as follows:

The rectangle R_1 is determined by the vertical lines through $x_{1,1}$ and $x_{1,7}$ and the horizontal lines through the sample points with smallest and largest ordinate, restricting ourselves to points which have abscissa values in the interior of the interval $[x_{1,1}, x_{1,7}]$. Similarly R_2 is determined by the vertical lines through

$x_{1,7}$ and $x_{1,13}$ and the horizontal lines through the sample points with largest and smallest ordinate, restricting ourselves to points with abscissa values in the interior of $[x_{1,7}, x_{1,13}]$. Finally R_3 is determined by the vertical lines through $x_{1,13}$ and $x_{1,19}$ and the horizontal lines through the sample points with largest and smallest ordinate, restricting ourselves to points whose abscissa values lie in the interior of $[x_{1,13}, x_{1,19}]$. The region T' consisting of the rectangles R_1, R_2 and R_3 has a much smaller area than the region T. As we will see later, the probability distribution of the statistic $\iint_{T'} f(x_1, x_2)dx_1\,dx_2$ is exactly the same as that of $\iint_{T} f(x_1, x_2)dx_1\,dx_2$. Thus the use of T' may be preferred to that of T.

We will consider tolerance regions T^* of the following general shape: Let $m_1, \cdots, m_k$ be k positive integers such that $1 \leq m_1, m_k \leq n$ and $m_{i+1} - m_i \geq 3$ where n is the size of the bivariate sample. Let V_i be the vertical line in the (x_1, x_2) plane given by the equation $x_1 = x_{1,m_i}$ $(i = 1, \cdots, k)$. The number of sample points which lie between the vertical lines V_i and V_{i+1} is obviously equal to $m_{i+1} - m_i - 1$. Through each point which lies between the vertical lines V_i and V_{i+1} we draw a horizontal line. In this way we obtain $m_{i+1} - m_i - 1$ horizontal lines $W_{i,1}, \cdots, W_{i,m_{i+1}-m_i-1}$ where the line $W_{i,j+1}$ is above the line $W_{i,j}$. Denote by R_{ij} $(i = 1, \cdots, k-1; j = 1, \cdots, m_{i+1} - m_i - 2)$ the rectangle determined by the lines V_i, V_{i+1}, $W_{i,j}$, $W_{i,j+1}$. Let T^* be a region composed of s different rectangles R_{ij}. The regions T and T' in the example illustrated in figure 1 are special cases of the type of regions T^* as described above. For the region T we have $k = 2$, $m_1 = 3$, $m_2 = 17$, $s = 12$, and for the region T' we have $k = 4$, $m_1 = 1$, $m_2 = 7$, $m_3 = 13$, $m_4 = 19$ and $s = 12$.

Let Q^* be given by $\iint_{T^*} f(x_1, x_2)dx_1\,dx_2$. We will prove that the probability density function of Q^* is given by

$$\frac{\Gamma(n+1)}{\Gamma(s)\Gamma(n-s+1)}\,Q^{*s-1}(1-Q^*)^{n-s}\,dQ^*. \tag{27}$$

Let $f_i(x_2)\,dx_2$ be the conditional distribution of x_2 under the restriction that $x_{1,m_i} < x_1 < x_{1,m_{i+1}}$. Thus, we have

$$f_i(x_2) = \frac{\int_{x_{1,m_i}}^{x_{1,m_{i+1}}} f(x_1, x_2)\,dx_1}{\int_{-\infty}^{+\infty}\int_{x_{1,m_i}}^{x_{1,m_{i+1}}} f(x_1, x_2)\,dx_1\,dx_2}. \tag{28}$$

Denote by $\varphi(x_1)\,dx_1$ the marginal distribution of x_1, i.e.

$$\varphi(x_1) = \int_{-\infty}^{+\infty} f(x_1, x_2)\,dx_2$$

Let

$$P_i^* = \int_{x_{1,m_i}}^{x_{1,m_i+1}} \varphi(x_1)\, dx_1 \qquad (i = 1, \cdots, k-1) \tag{29}$$

and

$$\overline{P}_i^* = \sum_j \int_{a_{ij}}^{b_{ij}} f_i(x_2)\, dx_2 \qquad (i = 1, \cdots, k-1) \tag{30}$$

where a_{ij} is the ordinate of the lower corners and b_{ij} is the ordinate of the upper corners of the rectangle R_{ij} and the summation is to be taken over all values of j for which R_{ij} is included in T^*. It is clear that

$$Q^* = P_1^* \overline{P}_1^* + \cdots + P_{k-1}^* \overline{P}_{k-1}^* . \tag{31}$$

Let y be any random variable which has a continuous probability density function, say $\psi(y)\, dy$. Furthermore let $y_1, \cdots, y_n$ be n independent observations on y. Let $\psi_i(y)\, dy$ be the conditional density function of y under the condition that y is restricted to the interval $[y_{m_i}, y_{m_i+1}]$. Let

$$P' = \sum_{i,j} \int_{y_{m_i+j}}^{y_{m_i+j+1}} \psi(y)\, dy \tag{32}$$

where the summation is taken over all pairs i, j for which R_{ij} is contained in T^*. Let

$$P_i' = \int_{y_{m_i}}^{y_{m_i+1}} \psi(y)\, dy, \qquad \text{and}$$

$$\overline{P}_i' = \sum_j \int_{y_{m_i+j}}^{y_{m_i+j+1}} \psi_i(y)\, dy,$$

where the summation is to be taken over all values j for which R_{ij} is contained in T^*. We obviously have

$$P' = P_1' \overline{P}_1' + \cdots + P_{k-1}' \overline{P}_{k-1}' . \tag{33}$$

It is easy to verify that (i) the joint distribution of $P_1', \cdots, P_{k-1}'$ is the same as the joint distribution of $P_1^*, \cdots, P_{k-1}^*$; (ii) the distribution of $\overline{P}_i'$ is the same as that of $\overline{P}_i^*$ $(i = 1, \cdots, k-1)$; (iii) the variates $\overline{P}_1', \cdots, \overline{P}_{k-1}'$ are independent of each other and also of $P_1', \cdots, P_{k-1}'$; (iv) the variates $\overline{P}_1^*, \cdots, \overline{P}_{k-1}^*$ are independent of each other and also of $P_1^*, \cdots, P_{k-1}^*$. Hence it follows from (31) and (33) that the distribution of Q^* is the same as that of P'. Now we will derive the distribution of P'. The expression P' can be written in the following form:

$$P' = \sum_{i=1}^{l} \int_{y_{r_i}}^{y_{s_i}} \psi(y)\, dy, \tag{34}$$

where $r_1, s_1, \cdots, r_l, s_l$ are some positive integers for which $1 \leq r_1 < s_1 < r_2 < s_2 < \cdots < r_l < s_l \leq n$. Let

$$P'' = \sum_{i=1}^{l-1} \int_{y_{r_i}}^{y_{s_i}} \psi(y)\, dy + \int_{y_{s_{l-1}}}^{y_{s_{l-1}+s_l-r_l}} \psi(y)\, dy$$

$$= \sum_{i=1}^{l-2} \int_{y_{r_i}}^{y_{s_i}} \psi(y)\, dy + \int_{y_{r_{l-1}}}^{y_{s_{l-1}+s_l-r_l}} \psi(y)\, dy. \tag{35}$$

For any fixed value $y_{s_{l-1}}$ denote by $\psi_1(y)$ the conditional probability density of y under the restriction that $y < y_{s_{l-1}}$ and by $\psi_2(y)$ the conditional distribution of y under the restriction that $y > y_{s_{l-1}}$. Let

$$P = \int_{-\infty}^{y_{s_{l-1}}} \psi(y)\, dy \qquad P_1 = \sum_{i=1}^{l-1} \int_{y_{r_i}}^{y_{s_i}} \psi_1(y)\, dy;$$

$$P_2 = \int_{y_{r_l}}^{y_{s_l}} \psi_2(y)\, dy \quad \text{and} \quad P_3 = \int_{y_{s_{l-1}}}^{y_{s_{l-1}+s_l-r_l}} \psi_2(y)\, dy.$$

Then it follows from (34) and (35) that

$$P' = PP_1 + (1 - P)P_2, \qquad P'' = PP_1 + (1 - P)P_3. \tag{36}$$

For calculating the distributions of P_2 and P_3 we may consider the variates $y_{s_{l-1}+1}, \cdots, y_n$ as $n - s_{l-1}$ independent observations drawn from a population which has the distribution $\psi_2(y)\, dy$. Hence, the distribution of P_2 can be derived from (13) and it is easy to verify that the distribution of P_3 is the same as that of P_2. It is clear that P_2 is independent of P and P_1. Similarly P_3 is independent of P and P_1. Hence, because of (36) the distribution of P' must be the same as that of P''.

In the same way we find that the distribution of P'' is the same as the distribution of

$$P''' = \sum_{i=1}^{l-3} \int_{y_{r_i}}^{y_{s_i}} \psi(y)\, dy + \int_{y_{r_{l-2}}}^{y_{s_{l-2}+s_{l-1}+s_l-r_l-r_{l-1}}} \psi(y)\, dy$$

Thus, by induction we see that the distribution of P' is the same as the distribution of the statistic $P_0 = \int_{y_{r_1}}^{y_{r_1+s}} \psi(y)\, dy$ where $s = \sum_{i=1}^{l} (s_i - r_i)$. From (13) it follows that the distribution of P_0 is given by

$$\frac{\Gamma(n+1)}{\Gamma(s)\Gamma(n-s+1)} P_0^{s-1}(1 - P_0)^{n-s}\, dP_0.$$

Hence, we have proved that the distribution of Q^* is given by (27).

5. Summary of the results and numerical illustrations. I shall give here a summary of the results obtained and a few illustrative examples. The multivariate case being a straightforward extension of the bivariate case, I shall discuss merely the latter. Consider a pair of random variables x and y. Denote by $f(x, y)dx\,dy$ the joint probability density function of x and y and suppose that nothing is known about $f(x, y)$ except that it is continuous. A sample of n pairs of independent observations $(x_1, y_1), \cdots, (x_n, y_n)$ is drawn from this bivariate population. The sample can be represented by n points $p_1, \cdots, p_n$ in the plane (x, y), p_i being the point with the coordinates x_i and y_i. In section 2 we have dealt with the problem of finding a rectangle T in the plane (x, y), called tolerance region, such that we can state with high probability, say with probability .98 or .99, that the proportion Q of the bivariate universe included in the rectangle T is not less than a given number b, say not less than .98 or .99. The rectangle T is constructed as follows: Suppose that the points $p_1, \cdots, p_n$ are arranged in order of increasing magnitude of their abscissa values, i.e. $x_1 < x_2 < \cdots < x_n$. We draw a vertical line V_{r_1} through the point p_{r_1} and a vertical line V_{s_1} through p_{s_1} where r_1 and s_1 are positive integers such that $1 \leq r_1$, $r_1 \leq s_1 - 3$ and $s_1 \leq n$. We consider the set S consisting of the points $p_{r_1+1}, \cdots, p_{s_1-1}$ which lie between the vertical lines V_{r_1} and V_{s_1}. We draw a horizontal line H_{r_2} through the point of S which has the r_2-th smallest ordinate in S. Finally a horizontal line H_{s_2} is drawn through the point of S which has the s_2-th smallest ordinate in S. The values r_2 and s_2 are positive integers for which $r_2 < s_2$. The tolerance region T is the rectangle determined by the lines V_{r_1}, V_{s_1}, H_{r_2} and H_{s_2}. The probability p that at least the porportion $b(0 < b < 1)$ of the universe is included in T is given by

$$(37) \qquad p = \int_b^1 \frac{\Gamma(n+1)}{\Gamma(s_2 - r_2)\Gamma(n - s_2 + r_2 + 1)} Q^{s_2 - r_2 - 1}(1 - Q)^{n - s_2 + r_2}\, dP.$$

It is known that if a random variable $v(0 \leq v \leq 1)$ has the distribution

$$\frac{\Gamma(c+d)}{\Gamma(c)\Gamma(d)} v^{c-1}(1 - v)^{d-1}\, dv,$$

and $2c$ and $2d$ are positive integers, then $\frac{2c}{2d}\frac{1 - v}{v}$ has the F-distribution (analysis of variance distribution) with $2d$ and $2c$ degrees of freedom. Thus,

$$(39) \qquad \frac{2(s_2 - r_2)}{2(n - s_2 + r_2 + 1)} \frac{1 - Q}{Q} = F$$

has the F-distribution with $2(n - s_2 + r_2 + 1)$ and $2(s_2 - r_2)$ degrees of freedom. From (37) it follows that p is equal to the probability that

$$F \leq \frac{2(s_2 - r_2)}{2(n - s_2 + r_2 + 1)} \frac{1 - b}{b}$$

where F has the analysis of variance distribution with $2(n - s_2 + r_2 + 1)$ and $2(s_2 - r_2)$ degrees of freedom. For the case $r_1 = 1$, $s_1 = n$, $r_2 = 1$ and $s_2 =$

$n - 2$, the following table gives the value of the sample size n which is necessary for having the probability p that at least the proportion b of the universe is included in the tolerance rectangle T.

	$b = .97$	$b = .975$	$b = .98$	$b = .985$	$b = .99$
$p = .99$	332	398	499	668	1001
$p = .95$	256	309	385	515	771

Thus, if we want the probability to be .99 that the tolerance region will include at least 98 per cent of the universe, the sample size must be 499.

In section 4 tolerance regions are considered which are composed of several rectangles. Such a tolerance region may be more favorable than a single rectangle if x and y are highly correlated. As an illustration we consider tolerance regions T^* constructed as follows: Suppose that n is divisible by 4 and the sample points $p_1, \ldots, p_n$ are arranged in order of increasing magnitude of their abscissa values. We draw the vertical lines V_0, V_1, V_2, V_3 and V_4 through the points p_1, $p_{n/4}$, $p_{n/2}$, $p_{3n/4}$ and p_n. Let $R_i (i = 1, 2, 3, 4)$ be the rectangle determined by the vertical lines V_{i-1} and V_i and the horizontal lines H_i and H_i' where H_i and H_i' are defined as follows: consider only the points which lie between the two vertical lines V_{i-1} and V_i (points on the vertical lines are excluded). From these select the point with the smallest and the point with the largest ordinate. The lines H_i and H_i' are the horizontal lines which go through these two points respectively. The tolerance region T^* is composed of the four rectangles R_1, R_2, R_3 and R_4. The number of rectangles R_{ij} (defined in section 4) included in T^* is equal to $s = n - 9$. Thus, according to the results of section 4 the probability distribution of the proportion Q^* of the universe included in the region T^* is given by

$$\frac{\Gamma(n+1)}{\Gamma(n-9)\Gamma(10)} (Q^*)^{n-8} (1 - Q^*)^9 \, dQ^*.$$

Numerical calculations show that for $n = 1000$ the probability is .99 that at least 98.1 per cent of the universe will be included in the tolerance region T^*.

ASYMPTOTIC FORMULAS FOR SIGNIFICANCE LEVELS OF CERTAIN DISTRIBUTIONS

By Alfred M. Peiser

Cornell University

1. Introduction. The purpose of this paper is to derive asymptotic formulas for the significance levels, or per cent points, of certain well-known statistical distributions.[1] Although we restrict ourselves here to two distributions, those of Chi-Square and of Student's t, it will be apparent that the methods used are applicable to many other distributions as well.

The following results are obtained. Let y_p be the p per cent point of the normal distribution, that is, the distribution defined by

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}v^2}\, dv, \tag{1.1}$$

so that

$$\Phi(y_p) = 1 - p. \tag{1.2}$$

If $\chi^2_{p,n}$ and $t_{p,n}$ denote the p per cent points of the Chi-Square and Student's t distributions with n degrees of freedom respectively, then

$$\chi^2_{p,n} = n + y_p\sqrt{2n} + \frac{2}{3}(y_p^2 - 1) + \frac{y_p^3 - 7y_p}{9\sqrt{2n}} + o\left(\frac{1}{\sqrt{n}}\right), \tag{1.3}$$

and

$$t_{p,n} = y_p + \frac{y_p^3 + y_p}{4n} + o\left(\frac{1}{n}\right). \tag{1.4}$$

These formulas approximate the true values of $\chi^2_{p,n}$ and $t_{p,n}$ to a high degree of accuracy. Tables of comparative values for several values of p and n are given in Section 4.

We shall need the following theorem due to Cramér [3, p. 81; see also pp. 86–87].

Theorem 1: *Let $X_1, X_2, \cdots$ be a sequence of independent, identically distributed random variables having an absolutely continuous distribution function with mean value zero, dispersion σ and finite fifth absolute moment. Let $H_n(x)$ be the distribution function of $(X_1 + \cdots + X_n)/(\sigma\sqrt{n})$, and let $n^{-\frac{1}{2}(r-2)}\lambda_r$ denote the r-th semi-invariant of $H_n(x)$. Then*

$$\Phi(x) - H_n(x) = \frac{\lambda_3}{3!\sqrt{n}}\Phi^{(3)}(x) - \frac{\lambda_4}{4!\,n}\Phi^{(4)}(x) - \frac{10\lambda_3^2}{6!\,n}\Phi^{(6)}(x) + 0(n^{-3/2}). \tag{1.5}$$

---

[1] This problem was proposed to the author by J. H. Curtiss.

2. The Chi-Square distribution. A random variable X is said to be distributed according to Chi-Square with n degrees of freedom ($X = \chi_n^2$) if its distribution function is

$$F_n(x) = \begin{cases} \displaystyle\int_0^x \frac{t^{\frac{1}{2}n-1}e^{-\frac{1}{2}t}}{2^{\frac{n}{2}}\Gamma(\frac{n}{2})}\,dt, & x \geqq 0, \\ 0, & x < 0. \end{cases} \tag{2.1}$$

The variable $(\chi_n^2 - n)/\sqrt{2n}$ then has the distribution function

$$H_n(x) = F_n(n + x\sqrt{2n}). \tag{2.2}$$

If we write

$$\chi_{p,n}^2 = n + y_p\sqrt{2n} + a_n, \tag{2.3}$$

so that

$$F_n(\chi_{p,n}^2) = 1 - p, \tag{2.4}$$

and let $z_{pn} = y_p + a_n/\sqrt{2n}$, then $H_n(z_{pn}) = 1 - p$, and it follows from (1.1) and (1.2) that

$$\Phi(z_{pn}) - H_n(z_{pn}) = \Phi(z_{pn}) - \Phi(y_p) = \frac{1}{\sqrt{2\pi}}\frac{a_n}{\sqrt{2n}}e^{-\frac{1}{2}(y_p+\theta a_n/\sqrt{2n})^2}, \tag{2.5}$$

where $0 < \theta < 1$. Then by a theorem of Liapounoff's [3, p. 77],

$$\frac{|a_n|}{\sqrt{2n}}e^{-\frac{1}{2}(y_p+\theta a_n/\sqrt{2n})^2} \leq \frac{K\log n}{\sqrt{n}},$$

where K denotes a constant independent of n. But if $\lim_{n\to\infty} |a_n|/\sqrt{2n} = \infty$, then $\lim_{n\to\infty} H_n(z_{pn})$ is either 0 or 1. Hence $a_n = o(\sqrt{n})$.

Fisher [1, p. 81] has suggested the use of

$$\chi_{p,n}^2 \doteq \tfrac{1}{2}[y_p + \sqrt{2n-1}]^2.$$

A closer approximation,

$$\chi_{p,n}^2 \doteq n\left[1 - \frac{2}{9n} + y_p\sqrt{\frac{2}{9n}}\right]^3,$$

has been obtained by Wilson and Hilferty [2]. It is interesting to note that, according to (1.3), this last approximation is correct to terms of the zero-th order in n.

We apply Theorem 1 to the variables $X_j = (\chi_1^2 - 1)/\sqrt{2}$, $j = 1, 2, \cdots$. Then $\sigma = 1$, and, by the additive property of the Chi-Square distribution [3, p. 45], $H_n(x)$, the distribution function of the variable $(X_1 + \cdots + X_n)/\sqrt{n}$,

is related to $F_n(x)$ by (2.2). Thus, $\lambda_3 = 2\sqrt{2}$. It follows from (1.5) and (2.3) that

$$\lim_{n\to\infty} \sqrt{2n}\,[\Phi(z_{pn}) - H_n(z_{pn})] = \lim_{n\to\infty} \frac{2}{3\sqrt{2\pi}}\,(z_{pn}^2 - 1)e^{-\frac{1}{2}z_{pn}^2}$$

$$= \frac{2}{3\sqrt{2\pi}}\,(y_p^2 - 1)e^{-\frac{1}{2}y_p^2}, \tag{2.6}$$

since $a_n = o(\sqrt{n})$. Then by (2.5) and (2.6)

$$\lim_{n\to\infty} a_n = \tfrac{2}{3}(y_p^2 - 1).$$

According to (2.3) we may now write

$$\chi_{p,2n}^2 = 2n + 2y_p\sqrt{n} + 2r_p + 2b_n\,,$$

where

$$r_p = \tfrac{1}{3}(y_p^2 - 1), \tag{2.7}$$

and $b_n = o(1)$. A simple change of variables in (2.1) yields

$$F_{2n}(\chi_{p,2n}^2) = \int_{\frac{n+r_p}{\sqrt{n}}}^{y_p+\frac{b_n}{\sqrt{n}}} \frac{\sqrt{n}e^{-n}n^n}{\Gamma(n+1)}\,e^{-v\sqrt{n}-r_p}\left[1 + \frac{v}{\sqrt{n}} + \frac{r_p}{n}\right]^{n-1} dv. \tag{2.8}$$

If we let

$$J_n = \int_{y_p}^{y_p+\frac{b_n}{\sqrt{n}}} \frac{1}{\sqrt{2\pi}}\,e^{-\frac{1}{2}v^2}\,dv, \tag{2.9}$$

then

$$nJ_n = \frac{b_n\sqrt{n}}{\sqrt{2\pi}}\,e^{-\frac{1}{2}(y_p+\delta_n)^2}, \tag{2.10}$$

where $\delta_n = o(1)$. By (1.2) and (2.4),

$$J_n = \Phi\left(y_p + \frac{b_n}{\sqrt{n}}\right) - F_{2n}(\chi_{p,2n}^2). \tag{2.11}$$

Using Stirling's formula for $\Gamma(n+1)$ in (2.8), (2.11) becomes

$$J_n = \int_{y_p+\frac{b_n}{\sqrt{n}}}^{\infty} \frac{e^{-\frac{1}{2}v^2}}{\sqrt{2\pi}}\left[\exp\left\{-\frac{1}{12n} + 0(n^{-3})\right.\right.$$

$$\left.\left. + (n-1)\log\left(1 + \frac{v}{\sqrt{n}} + \frac{r_p}{n}\right) + \frac{v^2}{2} - v\sqrt{n} - r_p\right\} - 1\right] dv$$

$$= \int_{y+\frac{b_n}{\sqrt{n}}}^{\infty} \frac{e^{-\frac{1}{2}v^2}}{\sqrt{2\pi}}\,[e^{A_n} - 1]\,dv,$$

where $A_n = \frac{1}{n}\left(-\frac{1}{12} - r_p - \frac{r_p^2}{2} + \frac{v^2}{2} + v^2 r_p - \frac{v^4}{4}\right) + \frac{1}{\sqrt{n}}\left(\frac{v^3}{3} - v - vr_p\right) + f_n(v)$, $nf_n(v)$ being dominated by $P(|\,v\,|)$, where P is a polynomial in v independent of n, and $f_n(v) = 0(n^{-3/2})$. Then

$$J_n = \int_{y_p+\frac{b_n}{\sqrt{n}}}^{\infty} \frac{e^{-\frac{1}{2}v^2}}{n\sqrt{2\pi}}\left[-\frac{1}{12} - r_p - \frac{r_p^2}{2} + v^2\left(1 + 2r_p + \frac{r_p^2}{2}\right)\right.$$

$$\left. -v^4\left(\frac{7}{12} + \frac{r_p}{3}\right) + \frac{v^6}{18}\right] dv + \int_{y_p+\frac{b_n}{\sqrt{n}}}^{\infty} \frac{e^{-\frac{1}{2}v^2}}{\sqrt{2\pi n}}\left[\frac{v^3}{3} - v - vr_p\right] dv \tag{2.12}$$

$$+ \int_{y_p+\frac{b_n}{\sqrt{n}}}^{\infty} \frac{e^{-\frac{1}{2}v^2}}{\sqrt{2\pi}}\, g_n(v)\, dv + \int_{y_p+\frac{b_n}{\sqrt{n}}}^{\infty} \frac{e^{-\frac{1}{2}v^2}}{\sqrt{2\pi}}\left(\sum_{j=3}^{\infty} \frac{A_n^j}{j!}\right) dv,$$

where $g_n(v)$ has the same properties given above for $f_n(v)$. If we call these last integrals K_1, K_2, K_3 and K_4 respectively, we see that

$$\lim_{n\to\infty} nK_3 = \int_{y_p}^{\infty} \frac{e^{-\frac{1}{2}v^2}}{\sqrt{2\pi}} \lim_{n\to\infty} ng_n(v)\, dv = 0. \tag{2.13}$$

Also, since A_n^j, $j = 1, 2, \cdots$ is dominated by $P_j(|\,v\,|)$, $P_j(v)$ being a polynomial in v independent of n, we see that

$$\sum_{j=3}^{\infty} \int_{y_p+\frac{b_n}{\sqrt{n}}}^{\infty} \frac{e^{-\frac{1}{2}v^2}}{\sqrt{2\pi}} \frac{|\,A_n\,|^j}{j!}\, dv \leq \sum_{j=3}^{\infty} e^{-\frac{1}{2}\left(y_p+\frac{b_n}{\sqrt{n}}\right)^2} \frac{Q_j\left(y_p + \frac{b_n}{\sqrt{n}}\right)}{j!},$$

where Q_j is a polynomial. Since this last sum converges, we have

$$nK_4 = \sum_{j=3}^{\infty} \int_{y_p+\frac{n}{\sqrt{n}}}^{\infty} \frac{e^{-\frac{1}{2}v^2}}{\sqrt{2\pi}} \frac{nA_n^j}{j!}\, dv, \tag{2.14}$$

and by the uniform convergence of (2.14),

$$\lim_{n\to\infty} nK_4 = \sum_{j=3}^{\infty} \int_{y_p}^{\infty} \frac{e^{-\frac{1}{2}v^2}}{j!\sqrt{2\pi}} \lim_{n\to\infty} nA_n^j\, dv = 0, \tag{2.15}$$

since $A_n^j = 0(n^{-j/2})$.

Integrating by parts we obtain

$$nK_2 = \frac{1}{\sqrt{2\pi}}\, e^{-\frac{1}{2}\left(y_p+\frac{b_n}{\sqrt{n}}\right)^2}\left(\frac{2}{3}\, y_p b_n + \frac{b_n^2}{3\sqrt{n}}\right),$$

and since $b_n = o(1)$,

$$\lim_{n\to\infty} nK_2 = 0. \tag{2.16}$$

Further integration by parts and the use of (2.7) yields

$$\lim_{n\to\infty} nK_1 = \frac{e^{-\frac{1}{2}y_p^2}}{36\sqrt{2\pi}}(y_p^3 - 7y_p). \tag{2.17}$$

Then, by (2.10), (2.12), (2.13), (2.15), (2.16) and (2.17),

$$\lim_{n\to\infty} b_n\sqrt{n} = \frac{1}{36}(y_p^3 - 7y_p),$$

so that

$$\chi^2_{p,2n} = 2n + 2y_p\sqrt{n} + \frac{2}{3}(y_p^2 - 1) + \frac{y_p^3 - 7y_p}{18\sqrt{n}} + o\left(\frac{1}{\sqrt{n}}\right).$$

Equation (1.3) now follows at once.

3. Student's t. If the random variable Y_n has the distribution function $\Phi(x/\sqrt{n})$, then $t_n = Y_n/\chi_n$ is distributed according to Student's distribution for n degrees of freedom and has the distribution function

$$G_n(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{n\pi}} \frac{\Gamma[\frac{1}{2}(n+1)]}{\Gamma[\frac{1}{2}n]} \left(1 + \frac{t^2}{n}\right)^{-\frac{1}{2}(n+1)} dt.$$

If $\sigma = \sqrt{n/(n-2)}$, the variable t_n/σ then has the distribution function

$$H_n(x) = G_n(x\sigma). \tag{3.1}$$

If we write

$$t_{p,n} = y_p + a_n\,, \tag{3.2}$$

so that

$$G_n(t_{p,n}) = 1 - p, \tag{3.3}$$

and let $z_{pn} = t_{p,n}/\sigma$, then $H_n(z_{pn}) = 1 - p$, and it follows from (1.1) and (1.2) that

$$\begin{aligned}\Phi(z_{pn}) - H_n(z_{pn}) &= \Phi(z_{pn}) - \Phi(y_p) \\ &= \frac{1}{\sqrt{2\pi}}\left[y_p\left(\frac{1}{\sigma} - 1\right) + \frac{a_n}{\sigma}\right] e^{-\frac{1}{2}\left[y_p + \theta\left(y_p\left(\frac{1}{\sigma}-1\right)+a_n\right)\right]^2},\end{aligned} \tag{3.4}$$

where $0 < \theta < 1$. Then by Liapounoff's Theorem [3, p. 77],

$$\left|y_p\left(\frac{1}{\sigma} - 1\right) + \frac{a_n}{\sigma}\right| e^{-\frac{1}{2}\left[y_p + \theta\left(y_p\left(\frac{1}{\sigma}-1\right)+a_n\right)\right]^2} \leqslant \frac{K\log n}{\sqrt{n}},$$

where K denotes a constant independent of n. But if $\lim_{n\to\infty} |a_n| = \infty$, then $\lim_{n\to\infty} H_n(z_{pn})$ is either 0 or 1.

Hence $a_n = o(1)$.

We apply Theorem 1 to the variables $X_j = Y_n/\chi_n$, $j = 1, 2, \cdots$. Then $\sigma = \sqrt{n/(n-2)}$, and by the additive property of the normal distribution, $H_n(x)$, the distribution function of $(X_1 + \cdots + X_n)/(\sigma\sqrt{n})$, satisfies the relation (3.1). Thus $\lambda_3 = 0$ and $\lambda_4 = 6n/(n-4)$. It follows from (1.5) and (3.2) that

$$\lim_{n\to\infty} n[\Phi(z_{pn}) - H_n(z_{pn})] = \lim_{n\to\infty} \frac{n}{4(n-4)\sqrt{2\pi}} (z_{pn}^3 - 3z_{pn})e^{-\frac{1}{2}z_{pn}^2} \tag{3.5}$$
$$= \frac{1}{4\sqrt{2\pi}} (y_p^3 - 3y_p)e^{-\frac{1}{2}y_p^2},$$

since $a_n = o(1)$. By (3.4) and (3.5) we have

$$\lim_{n\to\infty} n\left[y_p\left(\frac{1}{\sigma} - 1\right) + \frac{a_n}{\sigma}\right] = \frac{y_p^3 - 3y_p}{4}.$$

But $\lim_{n\to\infty} n(1-\sigma)/\sigma = -1$, so that

$$\lim_{n\to\infty} na_n = \frac{y_p^3 + y_p}{4}.$$

Hence

$$a_n = \frac{y_p^3 + y_p}{4n} + o\left(\frac{1}{n}\right),$$

and equation (1.4) follows at once.

4. Tables. The following tables compare the true values of $\chi_{p,n}^2$ and $t_{p,n}$ with those obtained from (1.3) and (1.4). The true values [4], [5], (to three decimal places) are shown in *italics*.

TABLE 1

$\chi_{p,n}^2$

n \ p	.01	.05	.1	.5	.9
10	23.253 *23.209*	18.318 *18.307*	15.989 *15.987*	9.333 *9.342*	4.875 *4.865*
30	50.908 *50.892*	43.777 *43.773*	40.257 *40.256*	29.333 *29.336*	20.600 *20.599*
50	76.163 *76.154*	67.507 *67.505*	63.168 *63.167*	49.333 *49.335*	37.689 *37.689*
100	135.811 *135.807*	124.343 *124.342*	118.499 *118.498*	99.333 *99.334*	82.358 *82.358*

TABLE II

$t_{p,n}$

n \ p	.0125	.025	.05	.125	.25
10	2.579 *2.634*	2.197 *2.228*	1.797 *1.813*	1.212 *1.221*	0.700 *0.700*
30	2.354 *2.360*	2.039 *2.042*	1.696 *1.697*	1.171 *1.173*	0.683 *0.683*
60	2.298 *2.299*	2.000 *2.000*	1.670 *1.671*	1.161 *1.162*	0.679 *0.679*
120	2.270 *2.270*	1.980 *1.980*	1.658 *1.658*	1.156 *1.156*	0.677 *0.677*

REFERENCES

[1] R. A. Fisher, *Statistical methods for research workers*, Oliver and Boyd, 1925.

[2] E. B. Wilson and M. M. Hilferty, *National Academy of Sciences Proc.*, Vol. 17 (1931), p. 687.

[3] H. Cramér, *Random variables and probability distributions*, Cambridge University Press, 1937.

[4] M. Merrington, "Tables of percentage points of the t distribution," *Biometrika*, Vol. 32 (April 1942), p. 300.

[5] M. Merrington, "Numerical approximation to the percentage points of the χ^2 distribution," *Biometrika*, Vol. 32 (October 1941), p. 200.

GENERALIZATION OF POINCARÉ'S FORMULA IN THE THEORY OF PROBABILITY

By Kai Lai Chung

Tsing Hua University, Kunming, China

Let $p_{[m]}(1, \cdots, n)$, $(0 \leq m \leq n)$ denote the probability of the occurrence of exactly m events among the n arbitrary events $E_1, \cdots, E_n$; and $p_m(1, \cdots, n)$ $(1 \leq m \leq n)$ that of at least m. Let $p_{\nu_1 \cdots \nu_i}$ $(1 \leq i \leq n)$, where $(\nu_1 \cdots \nu_i)$ is a combination (without repetition) out of $(1, \cdots, n)$, denote the probability of the occurrence of $E_{\nu_1}, \cdots, E_{\nu_i}$ (without regard to the other events); and

$$S_0 = 1, \qquad S_i = \sum_{(\nu_1 \cdots \nu_i)} p_{\nu_1 \cdots \nu_i},$$

where the summation extends to all the combinations with i members out of $(1, \cdots, n)$.

Then Poincaré's formula may be written as follows:

$$p_{[0]}(1, \cdots, n) = \sum_{i=0}^{n} (-1)^i S_i.$$

An equivalent formula is:

$$p_1(1, \cdots, n) = \sum_{i=1}^{n} (-1)^{i-1} S_i.$$

The following conventions concerning the binomial coefficients are made:

$$\binom{0}{0} = 1, \qquad \binom{a}{b} = 0 \qquad \text{if } a < b \text{ or } b < 0.$$

Two generalizations, possibly due to de Mises, are

$$p_{[m]}(1, \cdots, n) = \sum_{i=m}^{n} (-1)^{(i-m)} \binom{i}{m} S_i;$$

$$p_m(1, \cdots, n) = \sum_{i=m}^{n} (-1)^{(i-m)} \binom{i-1}{m-1} S_i.$$

We notice that the probabilities appearing on the left-hand sides of these formulas are symmetrical with respect to the set of suffixes $(1, \cdots, n)$, and the sums on the right-hand sides are symmetrical in the same way.

As a natural generalization let us consider a probability which is symmetrical with respect to certain sub-sets of $(1, \cdots, n)$. We divide the n events into r sets:

$$E_{\nu_{11}}, \cdots, E_{\nu_{1n_1}}; E_{\nu_{21}}, \cdots, E_{\nu_{2n_2}}; \cdots; E_{\nu_{r1}}, \cdots, E_{\nu_{rn_r}};$$

where $n_1 + n_2 + \cdots + n_r = n$. And we ask for the probability that out of the first set of n_1 events exactly m_1 events occur; and out of the second set of n_2 events exactly m_2 events occur; and so on; and finally, out of the rth set of n_r events exactly m_r events occur. When this problem is solved the analogous problem

in which we replace some of the words "exactly" by "at least" can also be solved.

We denote the required probability by the left-hand side of the following *generalized Poincaré's formula*:

$$p_{[m_1],[m_2],\cdots,[m_r]}(\nu_{11}\cdots\nu_{1n_1};\nu_{21}\cdots\nu_{2n_2};\cdots;\nu_{r1}\cdots\nu_{rn_r})$$

$$(1)\qquad =\sum_{i_1=m_1}^{n_1}\sum_{i_2=m_2}^{n_2}\cdots\sum_{i_r=m_r}^{n_r}(-1)^{i_1+i_2+\cdots+i-m_1-m_2-\cdots-m_r}\cdot\binom{i_1}{m_1}\binom{i_2}{m_2}\cdots\binom{i_r}{m_r}S_{i_1,i_2,\cdots,i_r},$$

where

$$S_{i_1,i_2,\cdots,i_r}=\Sigma p_{\alpha_{11}\cdots\alpha_{1i_1}\alpha_{21}\cdots\alpha_{2i_2}\cdots\alpha_{r1}\cdots\alpha_{ri_r}},$$

the summation extending to all those combinations of α's such that for every $k=1,\cdots,r$, $(\alpha_{k1}\cdots\alpha_{ki_k})$ is a combination of i_k members out of $(\nu_{k1}\cdots\nu_{kn_k})$.

PROOF: Let $p_{[\nu_1\cdots\nu_i]}$ denote the probability of the occurrence of the events $E_{\nu_1},\cdots,E_{\nu_i}$ and *these only* out of $E_1,\cdots,E_n$. It is well-known and also easily seen that

$$p_{\alpha_1\cdots\alpha_a}=\sum_{b=0}^{n-a}\sum_{\beta}p_{[\alpha_1\cdots\alpha_a\beta_1\cdots\beta_b]}$$

where for a fixed b the second summation extends to all the combinations $(\beta_1\cdots\beta_b)$ of b members out of the "difference set" $(1,\cdots,n)-(\alpha_1\cdots\alpha_a)$.

Now let each p in each S on the right-hand side of (1) be decomposed into a sum of the $p_{[\nu_1\cdots\nu_i]}$'s in the last-written way. Consider a fixed

$$p_{[\mu_{11}\cdots\mu_{1j_1}\mu_{21}\cdots\mu_{2j_2}\cdots\mu_{r1}\cdots\mu_{rj_r}]},$$

where for every $k=1,\cdots,r$, $(\mu_k,\cdots,\mu_{kj_k})$ is a combination of j_k members out of $(\nu_{k1}\cdots\nu_{kn_k})$. It appears once in exactly $\binom{j_1}{i_1}\binom{j_2}{i_2}\cdots\binom{j_r}{i_r}$ terms in $S_{i_1,i_2,\cdots,i_r}$. Hence, its total contribution to the right-hand side of (1) is

$$\sum_{i_1=m_1}^{n_1}\sum_{i_2=m_2}^{n_2}\cdots\sum_{i_r=m_r}^{n_r}(-1)^{i_1+i_2+\cdots+i_r-m_1-m_2-\cdots-m_r}\cdot\binom{i_1}{m_1}\binom{i_2}{m_2}\cdots\binom{i_r}{m_r}\binom{j_1}{i_1}\binom{j_2}{i_2}\cdots\binom{j_r}{i_r}$$

$$=\prod_{k=1}^{r}\binom{m_k}{j_k}\sum_{i_k=m_k}^{n_k}(-1)^{i_k-m_k}\binom{j_k-m_k}{i_k-m_k}$$

$$=\prod_{k=1}^{r}\binom{j_k}{m_k}\sum_{h_k=0}^{j_k-m_k}(-1)^{h_k}\binom{j_k-m_k}{h_k}$$

$$=\prod_{k=1}^{r}\binom{j_k}{m_k}\cdot\begin{cases}1 & \text{if } j_k=m_k\\ 0 & \text{if otherwise}\end{cases}$$

$$=\begin{cases}1 & \text{if } j_k=m_k \text{ for every } k=1,\cdots,r\\ 0 & \text{otherwise.}\end{cases}$$

Therefore after the decompositions and the collecting of terms, the only p's remaining on the right-hand side of (1) are those in which for every $k = 1, \cdots, j$ we have $j_k = m_k$. Thus the right-hand side is reduced to

$$\Sigma p_{[\mu_{11}\cdots\mu_{1m_1}\mu_{21}\cdots\mu_{2m_2}\cdots\mu_{r1}\cdots\mu_{rm_r}]},$$

where the summation extends to all those combinations of μ's such that for every $k = 1, \cdots, r$, $(\mu_{k1} \cdots \mu_{km_k})$ is a combination of m_k members out of $(\nu_{k1} \cdots \nu_{kn_k})$. This is clearly equal to the left-hand side of (1). Q. E. D.

If we replace "exactly m_k" by "at least m_k" in the definition of the probability just considered, we replace in our notation the square-bracketed $[m_k]$ by an unbracketed m_k and we replace in our formula $\binom{i_k}{m_k}$ by $\binom{i_k - 1}{m_k - 1}$. This is proved as before, noting that we have

$$\sum_{i_k=m_k}^{n_k} (-1)^{i_k-m_k} \binom{i_k - 1}{m_k - 1}\binom{j_k}{i_k} = 1 \qquad \text{for } j_k = m_k, \cdots, n_k;$$

and identity which can be proved by induction on j_k.

A parallel generalization of Poincaré's formula is as follows: We ask for the probability that *either* out of the first set exactly m_1 events occur; *or* out of the second exactly m_2; $\cdots$; *or* finally, out of the rth set exactly m_r. That is, instead of repeated conjunctions we may consider repeated disjunctions. We denote the required probability by the left hand side of (2), then it is given in terms of the p's defined above in (1) by the right-hand side below:

$$(2)\quad W_{[m_1],[m_2],\cdots,[m_r]}(\nu_{11}, \cdots, \nu_{1n_1}; \nu_{21} \cdots \nu_{2n_2}; \cdots; \nu_{r1} \cdots \nu_{rn_r}) = p_{m_1,m_2,\cdots,m_r} - p_{m_1+1,m_2+1,\cdots,m_r+1}.$$

Other events symmetrical with respect to each of the sub-sets, in whose definition the words "and", "or", "exactly", "at least" appear arbitrarily, may be considered.

Lastly, we only mention that as a first application the formula (1) can be used to establish the formula

$$(n - k)\Sigma p_m(\nu_1 \cdots \nu_k) = (k + 1 - m)\Sigma p_m(\nu_1 \cdots \nu_{k+1}) + m\Sigma p_{m+1}(\nu_1 \cdots \nu_{k+1}),$$

first obtained by P. L. Hsu. For its significance we may refer to [1], and a continuation of that paper to be published shortly.

REFERENCE

[1] K. L. CHUNG, "On the probability of the occurrence of least m events among n arbitrary events," *Annals of Math. Stat.*, Vol. 12 (September 1941).

TABLES FOR TESTING RANDOMNESS OF GROUPING IN A SEQUENCE OF ALTERNATIVES

By Frieda S. Swed and C. Eisenhart

University of Wisconsin

When two different kinds of objects are arranged along a line they will form two or more distinct groups of like objects. Thus, in the arrangement: *aabbbab*, there are 3 *a*'s and 4 *b*'s forming 4 groups. In general, if there are m objects of one kind and n objects of another kind, there are in all

$$C_m^{m+n} = C_n^{m+n}$$

different arrangements possible. There will be no loss of generality if we assume that $m \leq n$.

If u is defined to be the number of distinct groups of like objects in any one arrangement, then the proportion of arrangements yielding u' or less groups is[1]

$$P\{u \leq u'\} = \frac{1}{C_m^{m+n}} \sum_{u=2}^{u'} f_u , \tag{1}$$

where

$$f_u = 2C_{k-1}^{m-1} \cdot C_{k-1}^{n-1}, \quad \text{when} \quad u = 2k, \quad \text{i.e. } u \text{ is even},$$

and

$$f_u = C_{k-1}^{m-1} \cdot C_{k-2}^{n-1} + C_{k-2}^{m-1} \cdot C_{k-1}^{n-1}, \quad \text{when} \quad u = 2k - 1, \quad \text{i.e. } u \text{ is odd},$$

for $k = 1, 2, \cdots, m + 1$. For example, if $m = n = 5$, then

$$P\{u = 2\} = \frac{f_2}{C_5^{10}} = \frac{2\{C_0^4 \cdot C_0^4\}}{C_5^{10}} = \frac{1}{126},$$

$$P\{u = 3\} = \frac{f_3}{C_5^{10}} = \frac{C_1^4 \cdot C_0^4 + C_0^4 \cdot C_1^4}{C_5^{10}} = \frac{8}{252}.$$

In a random arrangement (1) is the probability of $u \leq u'$.

The following tables have been prepared for use in testing data for randomness and for testing whether two samples are from the same population. Table I gives $P\{u \leq u'\}$ to 7 decimal places for $m \leq n \leq 20$ with a range of m from 2 to 20 inclusive whereas Table II gives correct values for u_ϵ for $\epsilon = .005, .01, .025, .05, .95, .975, .99$ and $.995$, where u_ϵ is the *largest* integer, u', for which $P\{u \leq u'\} \leq \epsilon$ when $\epsilon < .50$, and is the *smallest* integer, u', for which $P\{u \leq u'\} \geq \epsilon$ when $\epsilon > .50$. This table was obtained from Table I and covers the same

[1] W. L. Stevens, "Distribution of Groups in a Sequence of Alternatives" (*Annals of Eugenics*, Vol. IX, Part I (1939) pp. 10–17).

A. Wald and J. Wolfowitz, "On a Test Whether Two Samples are from the Same Population" (*Annals of Math. Stat.*, Vol. XI, No. 2, June (1940) pp. 147–162).

range of values of m and n. Table III gives values of u_ϵ for $m = n$ from 10 to 100. These values of u_ϵ were obtained by using the normal approximation given on page 151 of the Wald-Wolfowitz paper together with a correction for continuity not given in their article—this correction improved the approximation for small values of m and n. The values of u_ϵ for $m = n = 10$ through 20 are included in Table III although they can be obtained from Table II in order to check on the adequacy of the approximation. These values obtained with the approximation check with those of Table II except for the five underscored values. It appears that the approximation will be adequate in general for $m = n \geq 20$.

To illustrate the use of these tables to test randomness of an arrangement,[2] consider a case where one might suspect nonrandomness and, more specifically, expect too few groups. The arrangement of diseased and healthy plants in a row of a field might be such a case. For example, we might have the following plant arrangement:

$$H\,H\,H\,H\,H\,H\,H\,H\,H\,H\,D\,H\,D\,D\,D\,D\,H\,H\,H\,H\,H\,H\,H\,H\,H\,H,$$

where

$m = 5$, the number of diseased plants present,
$n = 20$, the number of healthy plants present,
$u' = 5$, the number of groups actually formed.

From Table I the probability associated with this arrangement is found to be .018,3512, which is the probability of $u \leq u'$. Since $P < .05$, we might elect to regard this as evidence of a tendency for the disease to be nonrandomly distributed among the plants in a row, knowing that if we look for an explanation of "clustering" whenever $P\{u \leq u'\} \leq .05$ we may expect to follow a false scent not more than one time in twenty in the long run.

When a control chart[3] suggests the presence of assignable causes of variation in a manufactured product flowing from a production line, an examination of various types of runs, e.g. the lengths and relative frequency of runs above and below the median of a sequence of values, may assist in diagnosing the nature of the cause. Dr. Walter A. Shewhart has given us such an instance: A sequence of observations dealing with corrosion suggested the presence of an assignable cause of variation. By the use of run charts an assignable cause of variation was tracked down in the measuring apparatus and an attempt was made to eliminate it. The original sequence examined with regard to runs above and below the median of the sequence exhibited an unexpectedly large number of runs of length 7 or more and as a result a significantly low value of

[2] W. L. Stevens (ibid).

[3] American Defense Emergency Standards Z1.1 and Z1.2 entitled "Guide for Quality Control" and "Control Chart Method of Analyzing Data" and American War Standard Z1.3 entitled "Control Chart Method of Controlling Quality During Production" (published by the American Standards Association, New York City).

u, and, if the assignable cause were not completely eliminated in the new design, we might expect too large a proportion of long runs above and below the median, and, hence, too few total runs. A sequence of 40 observations taken with the new measuring device yielded a total of 15 runs above and below the median of the sequence which is significantly fewer than would be expected to arise under a state of statistical control, since for $m = n = 20$, $P\{u \leq 15\} = .038$. This sequence is of special interest since the occurrence of too few runs suggested the assignable cause had not been entirely eliminated although no especially long runs, say of length 7 or more, occurred in this sequence, so that from the point of view of length of runs without regard to their number the assignable cause might have been judged to have been eliminated.

As an instance where too many groups would be the probable alternative to randomness consider the arrangement of occupied and unoccupied seats at a lunch counter about half an hour before the popular lunch hour begins. In such a case the critical region would be $u \geq u'$ and the appropriate probability would be $P = 1 - P\{u \leq u' - 1\}$. Such a situation was observed and yielded the following arrangement of empty and occupied seats along the lunch counter:

$$E\,O\,E\,E\,O\,E\,E\,E\,O\,E\,E\,E\,O\,E\,O\,E,$$

$$\begin{aligned} m &= 5, \\ n &= 11, \\ u' &= 11, \\ P &= 1 - .942{,}3077 = .057{,}6923; \end{aligned}$$

and though this probability is not quite significant, the arrangement observed has the maximum number of groups of empty and occupied seats for the m and n of the size observed since no two occupied seats are adjacent. However, if another customer had entered and sat either in the 5th empty seat from the left or in the 8th empty seat, the number of groups would have been increased by two and the situation would be:

$$\begin{aligned} m &= 6, \\ n &= 10, \\ u' &= 13, \\ P &= 1 - .989{,}5105 = .010{,}4895. \end{aligned}$$

This P value is significant, and for this assumed case, as well as for the actual case observed, the arrangement of E's and O's has the maximum number of groups of like objects. Certainly both of these cases exhibit too many groups to be considered random arrangements.

The use of these tables to test whether two samples constitute independent random samples from the same population[4] can be illustrated by using the data of Snedecor's Example 4.11 on page 75 of his *Statistical Methods* (3d edition)

[4] A. Wald and J. Wolfowitz (ibid) have pointed out that exceptionally small values of u' are to be regarded as evidence for rejecting this null hypothesis.

which gives daily gains in two lots of steer calves on two different rations. The daily rates of gain given for the two lots are:

I. 1.95, 2.17, 2.06, 2.11, 2.24, 2.52, 2.04, 1.95;
V. 1.82, 1.85, 1.87, 1.74, 2.04, 1.78, 1.76, 1.86.

Arranging these rates in order of magnitude, designating a calf on ration I by italics and one from V by ( ), we have

(1.74), (1.76), (1.78), (1.82), (1.85), (1.86), (1.87), *1.95*, *1.95*, (2.04), *2.04*, *2.06*, *2.11*, *2.17*, *2.24*, *2.52*.

Whence

$$m = 8,$$
$$n = 8,$$
$$u' = 4,$$
$$P = .008,8578.$$

Accordingly, at either the .05 or .01 level of significance rejection of the null hypothesis that the two samples constitute independent random samples from the same population is indicated.

For these data we note the fact that having two identical values, i.e. 2.04, in the two lots did not alter the number of groups regardless of whether they were recorded as (2.04), *2.04* or as *2.04*, (2.04). However, such duplications in general may be more bothersome, since they may yield different values of u' depending on the order in which they are considered. In such instances both possible orders should be considered.

The merit of this test is that it employs a minimum of assumptions—merely that the common population be continuous, and that the samples be drawn at random independently. Its principal defect is its lack of power. As a consequence gross disparity between the samples is generally required to render $u' \leq u_\epsilon$. Therefore, when additional assumptions are tenable, tests utilizing them should be employed.

Most of the computing and checking of these tables was done by Frieda S. Swed, Philip Ritz and Beatrice E. Kelley with some assistance from Jay Grodman, Edward Halamka and Mrs. Henry Wallman. Also, Duane Borst and Francis Cox helped with the typing and the proofing of the tables.

TABLE I

$P\{u \leq u'\}$

When m = n, the largest possible value of u′ is 2m; when m < n, the largest possible value of u′ is 2m + 1.

m = 2

n \ u′	2	3	4
2	.333,3333	.666,6667	1.
3	.200,0000	.500,0000	.900,0000
4	.133,3333	.400,0000	.800,0000
5	.095,2381	.333,3333	.714,2857
6	.071,4286	.285,7143	.642,8571
7	.055,5556	.250,0000	.583,3333
8	.044,4444	.222,2222	.533,3333
9	.036,3636	.200,0000	.490,9091
10	.030,3030	.181,8182	.454,5455
11	.025,6410	.166,6667	.423,0769
12	.021,9780	.153,8462	.395,6044
13	.019,0476	.142,8571	.371,4286
14	.016,6667	.133,3333	.350,0000
15	.014,7059	.125,0000	.330,8824
16	.013,0719	.117,6471	.313,7255
17	.011,6959	.111,1111	.298,2456
18	.010,5263	.105,2632	.284,2105
19	.009,5238	.100,0000	.271,4286
20	.008,6580	.095,2381	.259,7403

m = 3

n \ u′	2	3	4	5	6
3	.100,0000	.300,0000	.700,0000	.900,0000	1.
4	.057,1429	.200,0000	.542,8571	.800,0000	.971,4286
5	.035,7143	.142,8571	.428,5714	.714,2857	.928,5714
6	.023,8095	.107,1429	.345,2381	.642,8571	.880,9524
7	.016,6667	.083,3333	.283,3333	.583,3333	.833,3333
8	.012,1212	.066,6667	.236,3636	.533,3333	.787,8788
9	.009,0909	.054,5455	.200,0000	.490,9091	.745,4545
10	.006,9930	.045,4545	.171,3287	.454,5455	.706,2937
11	.005,4945	.038,4615	.148,3516	.423,0769	.670,3297
12	.004,3956	.032,9670	.129,6703	.395,6044	.637,3626
13	.003,5714	.028,5714	.114,2857	.371,4286	.607,1429
14	.002,9412	.025,0000	.101,4706	.350,0000	.579,4118
15	.002,4510	.022,0588	.090,6863	.330,8824	.553,9216
16	.002,0640	.019,6078	.081,5273	.313,7255	.530,4438
17	.001,7544	.017,5439	.073,6842	.298,2456	.508,7719
18	.001,5038	.015,7895	.066,9173	.284,2105	.488,7218
19	.001,2987	.014,2857	.061,0390	.271,4286	.470,1299
20	.001,1293	.012,9870	.055,9006	.259,7403	.452,8515

m = 4

n \ u′	2	3	4	5	6
4	.028,5714	.114,2857	.371,4286	.628,5714	.885,7143
5	.015,8730	.071,4286	.261,9048	.500,0000	.785,7143
6	.009,5238	.047,6190	.190,4762	.404,7619	.690,4762
7	.006,0606	.033,3333	.142,4242	.333,3333	.606,0606
8	.004,0404	.024,2424	.109,0909	.278,7879	.533,3333
9	.002,7972	.018,1818	.085,3147	.236,3636	.471,3287
10	.001,9980	.013,9860	.067,9321	.202,7972	.418,5814
11	.001,4652	.010,9890	.054,9451	.175,8242	.373,6264
12	.001,0989	.008,7912	.045,0549	.153,8462	.335,1648
13	.000,8403	.007,1429	.037,3950	.135,7143	.302,1008
14	.000,6536	.005,8824	.031,3725	.120,5882	.273,5294
15	.000,5160	.004,9020	.026,5736	.107,4431	.248,7100
16	.000,4128	.004,1280	.022,7038	.097,0072	.227,0382
17	.000,3342	.003,5088	.019,5489	.087,7193	.208,0200
18	.000,2734	.003,0075	.016,9515	.079,6992	.191,2509
19	.000,2259	.002,5974	.014,7939	.072,7273	.176,3975
20	.000,1882	.002,2546	.012,9870	.066,6250	.163,1846

m = 4

n \ u′	7	8
4	.971,4286	1.
5	.928,5714	.992,0635
6	.880,9524	.976,1905
7	.833,3333	.954,5455
8	.787,8788	.929,2929
9	.745,4545	.902,0979
10	.706,2937	.874,1259
11	.670,3297	.846,1538
12	.637,3626	.818,6813
13	.607,1429	.792,0168
14	.579,4118	.766,3399
15	.553,9216	.741,7441
16	.530,4438	.718,2663
17	.508,7719	.695,9064
18	.488,7218	.674,6411
19	.470,1299	.654,4325
20	.452,8515	.635,2343

n \ u'	1	2	3	4	5
20	.001,1293	.012,9870	.055,9006	.259,7403	.452,8515

n \ u'	1	2	3	4	5
20	.452,8515	.635,2343			

TABLE I (*Continued*)

n = 5

n \ u'	2	3	4	5	6
5	.007,9365	.039,6825	.166,6667	.357,1429	.642,8571
6	.004,3290	.023,8095	.110,3896	.261,9048	.521,6450
7	.002,5253	.015,1515	.075,7576	.196,9697	.424,2424
8	.001,5540	.010,1010	.053,6131	.151,5152	.347,3193
9	.000,9990	.006,9930	.038,9610	.118,8811	.286,7133
10	.000,6660	.004,9950	.028,9710	.094,9051	.238,7612
11	.000,4579	.003,6630	.021,9780	.076,9231	.200,5494
12	.000,3232	.002,7472	.016,9683	.063,1868	.169,8449
13	.000,2334	.002,1008	.013,3053	.052,5210	.144,9580
14	.000,1720	.001,6340	.010,5779	.044,1176	.124,6130
15	.000,1290	.001,2900	.008,5139	.037,4097	.107,8431
16	.000,0983	.001,0320	.006,9291	.031,9917	.093,9112
17	.000,0759	.000,8354	.005,6960	.027,5689	.082,2511
18	.000,0594	.000,6835	.004,7252	.023,9234	.072,4241
19	.000,0471	.000,5617	.003,9526	.020,8922	.064,0881
20	.000,0376	.000,4705	.003,3314	.018,3512	.056,9735

n = 6

n \ u'	2	3	4	5	6
6	.002,1645	.012,9870	.067,0996	.175,3247	.391,7749
7	.001,1655	.007,5758	.042,5408	.121,2121	.296,0373
8	.000,6660	.004,6620	.027,9720	.086,2471	.226,1072
9	.000,3996	.002,9970	.018,9810	.062,9371	.174,8252
10	.000,2498	.001,9980	.013,2368	.046,9530	.136,8631
11	.000,1616	.001,3736	.009,4538	.035,7143	.108,4357
12	.000,1077	.000,9696	.006,8951	.027,6341	.086,8886
13	.000,0737	.000,7003	.005,1231	.021,7087	.070,3597
14	.000,0516	.000,5160	.003,8700	.017,2859	.057,5335
15	.000,0369	.000,3870	.002,9670	.013,9319	.047,4716
16	.000,0268	.000,2949	.002,3052	.011,3519	.039,4971
17	.000,0198	.000,2278	.001,8128	.009,3415	.033,1164
18	.000,0149	.000,1783	.001,4414	.007,7565	.027,9652
19	.000,0113	.000,1412	.001,1575	.006,4935	.023,7719
20	.000,0087	.000,1129	.000,9382	.005,4771	.020,3318

n = 5

n \ u'	7	8	9	10
5	.833,3333	.960,3175	.992,0635	1.
6	.738,0952	.911,2554	.976,1905	.997,8355
7	.651,5152	.853,5354	.954,5455	.992,4242
8	.575,7576	.793,3178	.929,2929	.983,6830
9	.510,4895	.734,2657	.902,0979	.972,0280
10	.454,5454	.678,3217	.874,1259	.958,0420
11	.406,5934	.626,3736	.846,1538	.942,3077
12	.365,3846	.578,7007	.818,6813	.925,3394
13	.329,8319	.535,2474	.792,0168	.907,5630
14	.299,0196	.495,7860	.766,3399	.889,3189
15	.272,1878	.460,0103	.741,7441	.870,8720
16	.248,7100	.427,5886	.718,2662	.852,4252
17	.228,0702	.398,1924	.695,9064	.834,1308
18	.209,8428	.371,5118	.674,6411	.816,1015
19	.193,6759	.347,2614	.654,4325	.798,4190
20	.179,2772	.325,1835	.635,2343	.781,1406

n = 6

n \ u'	7	8	9	10	11	12
6	.608,2251	.824,6753	.932,9004	.987,0130	.997,8355	1.
7	.500,0000	.733,1002	.876,7879	.966,2005	.992,4242	.999,4172
8	.412,5874	.645,6876	.820,5128	.937,0629	.983,6830	.997,6690
9	.342,6573	.566,4336	.762,2378	.902,0979	.972,0280	.994,4056
10	.286,7133	.496,5035	.706,2937	.863,6364	.958,0420	.989,5105
11	.241,7582	.435,6820	.653,8462	.823,5294	.942,3077	.983,0317
12	.205,3975	.383,1610	.605,3652	.783,1286	.925,3394	.975,1131
13	.175,7703	.337,9404	.560,9244	.743,3658	.907,5630	.965,9443
14	.151,4448	.299,0196	.520,3818	.704,3504	.889,3189	.955,7276
15	.131,3209	.265,4799	.483,4881	.667,9567	.870,8720	.944,6594
16	.114,5511	.236,5137	.449,9484	.632,8924	.852,4252	.932,9205
17	.100,4785	.211,4278	.419,4577	.599,7504	.834,1308	.920,6712
18	.088,5910	.189,6342	.391,7204	.568,5469	.816,1015	.908,0508
19	.078,4867	.170,6381	.366,4596	.539,2434	.798,4190	.895,1779
20	.069,8475	.154,0242	.343,4218	.511,7752	.781,1406	.882,1526

TABLE I (*Continued*)

n = 7

n \ u'	2	3	4	5	6
7	.000,5828	.004,0793	.025,0583	.077,5058	.208,6247
8	.000,3108	.002,3310	.015,3846	.051,2821	.149,1841
9	.000,1748	.001,3986	.009,7902	.034,9650	.105,3916
10	.000,1028	.000,8741	.006,4274	.024,4755	.080,0082
11	.000,0628	.000,5656	.004,3363	.017,5339	.059,9548
12	.000,0397	.000,3771	.002,9967	.012,8205	.045,5664
13	.000,0258	.000,2580	.002,1156	.009,5459	.035,0877
14	.000,0172	.000,1806	.001,5222	.007,2239	.027,3478
15	.000,0117	.000,1290	.001,1141	.005,5470	.021,5546
16	.000,0082	.000,0938	.000,8280	.004,3156	.017,1645
17	.000,0058	.000,0693	.000,6241	.003,3978	.013,7993
18	.000,0042	.000,0520	.000,4764	.002,7044	.011,1920
19	.000,0030	.000,0395	.000,3679	.002,1739	.009,1517
20	.000,0023	.000,0304	.000,2872	.001,7635	.007,5403

n = 7

n \ u'	12	13	14
7	.995,9207	.999,4172	1.
8	.987,8788	.997,6690	.999,8446
9	.974,8252	.994,4056	.999,3007
10	.957,1164	.989,5105	.998,1489
11	.935,5204	.983,0317	.996,2293
12	.910,9312	.975,1131	.993,4508
13	.884,2105	.965,9443	.989,7833
14	.856,1146	.955,7276	.985,2425
15	.827,2704	.944,6594	.979,8762
16	.798,1783	.932,9205	.973,7515
17	.769,2254	.920,6712	.966,9464
18	.740,7031	.908,0508	.959,5423
19	.712,8246	.895,1779	.951,6206
20	.685,7403	.882,1526	.943,2587

n = 7

n \ u'	7	8	9	10	11
7	.383,4499	.616,5501	.791,3753	.922,4942	.974,9417
8	.296,0373	.513,5975	.703,9627	.867,1329	.948,7179
9	.230,7692	.426,5734	.622,3776	.805,9441	.916,0839
10	.181,8182	.354,5866	.548,9510	.743,3155	.879,3706
11	.144,7964	.295,6259	.484,1629	.682,1267	.840,4977
12	.116,5158	.247,4994	.427,6018	.624,0772	.800,9050
13	.094,6852	.208,2043	.378,4830	.570,0464	.761,6099
14	.077,6574	.176,0406	.335,9133	.520,3818	.723,2972
15	.064,2415	.149,6153	.299,0196	.475,1032	.686,4035
16	.053,5698	.127,8079	.267,0044	.434,0402	.651,1868
17	.045,0038	.109,7242	.239,1651	.396,9212	.617,7796
18	.038,0695	.094,6536	.214,8949	.363,4283	.586,2284
19	.032,4111	.082,0310	.193,6759	.333,2320	.556,5217
20	.027,7592	.071,4064	.175,0684	.306,0099	.528,6105

n = 8

n \ u'	2	3	4	5	6
8	.000,1554	.001,2432	.008,6578	.031,7016	.100,2331
9	.000,0823	.000,6993	.005,3065	.020,2797	.068,6549
10	.000,0457	.000,4114	.003,2908	.013,3690	.047,9227
11	.000,0265	.000,2514	.002,1037	.009,0498	.034,0557
12	.000,0159	.000,1588	.001,3813	.006,2713	.024,6090
13	.000,0098	.000,1032	.000,9288	.004,4376	.018,0599
14	.000,0063	.000,0688	.000,6380	.003,1992	.013,4440
15	.000,0041	.000,0469	.000,4467	.002,3454	.010,1404
16	.000,0027	.000,0326	.000,3182	.001,7458	.007,7420
17	.000,0018	.000,0231	.000,2302	.001,3175	.005,9774
18	.000,0013	.000,0166	.000,1690	.001,0069	.004,6631
19	.000,0009	.000,0122	.000,1257	.000,7784	.003,6728
20	.000,0006	.000,0090	.000,0946	.000,6081	.002,9185

TABLE I (*Continued*)

$n = 8$

n \ u'	7	8	9	10	11
8	.214,4522	.404,8174	.595,1826	.785,5478	.899,7669
9	.157,3427	.318,5932	.500,0000	.701,5631	.842,6573
10	.117,0300	.251,4055	.419,3747	.620,9379	.782,1884
11	.088,2353	.199,3729	.352,1870	.546,6778	.721,7195
12	.067,3970	.159,0855	.296,6182	.479,9952	.663,3722
13	.052,1156	.127,7950	.250,7740	.421,0526	.608 3591
14	.040,7637	.103,3712	.212,9343	.369,4530	.557,2755
15	.032,2263	.084,1930	.181,6305	.324,5390	.510,3199
16	.025,7305	.069,0369	.155,6472	.285,5639	.467,4474
17	.020,7337	.056,9771	.133,9944	.251,7856	.428,4724
18	.016,8504	.047,3188	.115,8727	.222,5120	.393,1350
19	.013,8036	.039,5324	.100,6385	.197,1217	.361,1432
20	.011,3915	.033,2151	.087,7741	.175,0684	.332,1982

$n = 9$

n \ u'	2	3	4	5	6
9	.000,0411	.000,3702	.003,0029	.012,2172	.044,4673
10	.000,0217	.000,2057	.001,7645	.007,6100	.029,4334
11	.000,0119	.000,1191	.001,0717	.004,8821	.019,8857
12	.000,0068	.000,0714	.000,6702	.003,2151	.013,6937
13	.000,0040	.000,0442	.000,4302	.002,1672	.009,5975
14	.000,0024	.000,0281	.000,2827	.001,4917	.006,8368
15	.000,0015	.000,0184	.000,1897	.001,0463	.004,9438
16	.000,0010	.000,0122	.000,1297	.000,7465	.003,6246
17	.000,0006	.000,0083	.000,0903	.000,5409	.002,6916
18	.000,0004	.000,0058	.000,0638	.000,3975	.002,0225
19	.000,0003	.000,0041	.000,0458	.000,2959	.001,5364
20	.000,0002	.000,0029	.000,0333	.000,2230	.001,1791

$n = 8$

n \ u'	12	13	14	15	16
8	.968,2984	.991,1422	.998,7568	.999,8446	1.
9	.939,4076	.979,7203	.995,8453	.999,3007	.999,9589
10	.903,1263	.963,5952	.990,4703	.998,1489	.999,7943
11	.861,7528	.943,4389	.982,3371	.996,2293	.999,4046
12	.817,4089	.920,1000	.971,4456	.993,4508	.998,6902
13	.771,8266	.894,4272	.957,9979	.989,7833	.997,5674
14	.726,3158	.867,1827	.942,3117	.985,2425	.995,9752
15	.681,8100	.839,0093	.924,7543	.979,8762	.993,8754
16	.638,9375	.810,4276	.905,6999	.973,7515	.991,2505
17	.598,0917	.781,8459	.885,5022	.966,9464	.988,1007
18	.559,4924	.753,5760	.864,4810	.959,5423	.984,4394
19	.523,2350	.725,8498	.842,9161	.951,6206	.980,2899
20	.489,3281	.698,8345	.821,0466	.943,2587	.975,6823

$n = 9$

n \ u'	7	8	9	10	11
9	.108,9675	.237,9679	.399,2184	.600,7816	.762,0321
10	.076,7174	.178,5598	.318,5932	.509,5477	.681,4068
11	.054,8940	.134,9131	.254,9417	.429,9833	.605,0250
12	.039,8905	.102,7626	.204,0297	.362,1100	.535,0083
13	.029,4118	.078,9474	.165,6347	.304,9536	.472,1362
14	.021,9814	.061,1792	.134,6749	.257,1679	.416,4087
15	.016,6363	.047,8163	.110,1763	.217,3577	.367,4115
16	.012,7388	.037,6828	.090,6888	.184,2289	.324,5330
17	.009,8606	.029,9339	.075,0988	.156,6466	.287,1229
18	.007,7099	.023,9597	.062,5530	.133,6459	.254,5038
19	.006,0849	.019,3169	.052,3969	.114,4218	.226,0667
20	.004,8444	.015,6810	.044,1269	.098,3096	.201,2567

TABLE I (*Continued*)

$n = 9$

$n-u'$	12	13	14	15	16
9	.891,0325	.955,5327	.987,7828	.996,9971	.999,6298
10	.834,1705	.923,2826	.974,2038	.992,3900	.998,6252
11	.773,0650	.885,0917	.955,1084	.985,1155	.996,5468
12	.711,0502	.843,0817	.931,1026	.975,1131	.993,0766
13	.650,4644	.799,0712	.903,0960	.962,5387	.988,0141
14	.592,7985	.754,4892	.872,0824	.947,6780	.981,2761
15	.538,9016	.710,3917	.839,0093	.930,8790	.972,8766
16	.489,1695	.667,5192	.804,7113	.912,5050	.962,9021
17	.443,6946	.626,3616	.769,8856	.892,9062	.951,4874
18	.402,3771	.587,2186	.735,0919	.872,4027	.938,7948
19	.365,0025	.550,2503	.700,7642	.851,2780	.924,9991
20	.331,2952	.515,5164	.667,2279	.829,7760	.910,2760

$n = 9$

$n-u'$	17	18
9	.999,9589	1..
10	.999,7943	.999,9892
11	.999,4046	.999,9405
12	.998,6902	.999,8129
13	.997,5674	.999,5577
14	.995,9752	.999,1251
15	.993,8754	.998,4688
16	.991,2505	.997,5501
17	.988,1007	.996,3387
18	.984,4394	.994,8131
19	.980,2899	.992,9607
20	.975,6823	.990,7760

$n = 10$

$n-u'$	2	3	4	5	6
10	.000,0108	.000,1083	.000,9851	.004,4924	.018,5217
11	.000,0057	.000,0595	.000,5699	.002,7387	.011,9246
12	.000,0031	.000,0340	.000,3402	.001,7181	.007,8420
13	.000,0017	.000,0201	.000,2089	.001,1057	.005,2593
14	.000,0010	.000,0122	.000,1315	.000,7281	.003,5916
15	.000,0006	.000,0076	.000,0847	.000,4895	.002,4939
16	.000,0004	.000,0049	.000,0557	.000,3353	.001,7586
17	.000,0002	.000,0032	.000,0373	.000,2336	.001,2578
18	.000,0002	.000,0021	.000,0255	.000,1654	.000,9115
19	.000,0001	.000,0014	.000,0176	.000,1187	.000,6687
20	.000,0001	.000,0010	.000,0124	.000,0864	.000,4962

$n = 10$

$n-u'$	7	8	9	10	11
10	.051,2668	.127,6386	.242,2113	.414,0704	.585,9296
11	.034,8893	.092,0457	.184,9250	.334,9607	.500,0000
12	.024,1724	.067,0398	.142,0576	.270,6597	.424,9821
13	.017,0279	.049,3337	.109,9071	.218,9393	.360,6811
14	.012,1820	.036,6806	.085,6777	.177,5474	.306,1650
15	.008,8413	.027,5493	.067,3038	.144,4744	.260,2302
16	.006,5028	.020,8935	.053,2728	.118,0313	.221,6449
17	.004,8423	.015,9941	.042,4797	.096,8449	.189,2657
18	.003,6475	.012,3527	.034,1158	.079,8184	.162,0831
19	.002,7769	.009,6211	.027,5869	.066,0851	.139,2318
20	.002,1353	.007,5536	.022,4538	.054,9634	.119,9827

$n = 10$

$n-u'$	12	13	14	15	16
10	.757,7887	.872,3614	.948,7432	.981,4783	.995,5076
11	.680,0429	.815,0750	.915,0988	.965,1107	.989,6064
12	.605,0250	.755,0607	.875,0893	.943,6771	.980,4205
13	.535,1326	.695,0464	.830,7309	.917,9567	.967,7999
14	.471,5305	.636,8959	.783,8875	.888,8814	.951,8778
15	.414,5713	.581,7741	.736,1152	.857,3832	.932,9789
16	.364,1136	.530,3271	.688,6257	.824,3101	.911,5359
17	.319,7420	.482,8375	.642,3087	.790,3890	.888,0244
18	.280,9098	.439,3454	.597,7810	.756,2166	.862,9182
19	.247,0268	.399,7365	.555,4405	.722,2661	.836,6609
20	.217,5115	.363,8048	.515,5164	.688,9010	.809,6510

TABLE I (*Continued*)

$n = 10$

n' \ u'	17	18	19	20
10	.999,0149	.999,8917	.999,9892	1.
11	.997,2613	.999,5577	.999,9405	.999,9972
12	.994,1993	.998,7922	.999,8129	.999,9830
13	.989,6064	.997,3944	.999,5577	.999,9423
14	.983,3760	.995,1878	.999,1251	.999,8542
15	.975,5014	.992,0380	.998,4688	.999,6938
16	.966,0520	.987,8584	.997,5501	.999,4346
17	.955,1487	.982,6087	.996,3387	.999,0508
18	.942,9443	.976,2885	.994,8131	.998,5180
19	.929,6066	.968,9298	.992,9607	.997,8154
20	.915,3073	.960,5885	.990,7760	.996,9253

$n = 11$

n' \ u'	12	13	14	15	16
11	.590,0214	.740,0572	.865,0869	.936,5325	.977,3586
12	.507,1756	.665,0393	.808,5517	.900,8097	.959,3862
13	.433,4692	.593,2831	.748,7549	.859,8062	.935,9556
14	.369,0039	.526,6523	.688,3430	.815,3856	.907,7803
15	.313,6559	.466,0183	.629,2637	.769,1883	.875,7976
16	.266,4962	.411,6032	.572,8332	.722,5468	.841,0015
17	.226,5446	.363,2342	.519,8576	.676,4811	.804,3369
18	.192,8141	.320,5187	.470,7594	.631,7315	.766,6415
19	.164,3540	.282,9585	.425,6872	.588,8056	.728,6214
20	.140,4323	.250,0211	.384,6040	.548,0260	.690,8486

$n = 11$

n' \ u'	2	3	4	5	6
11	.000,0028	.000,0312	.000,3147	.001,5905	.007,3317
12	.000,0015	.000,0170	.000,1797	.000,9526	.004,6136
13	.000,0008	.000,0096	.000,1058	.000,5865	.002,9662
14	.000,0004	.000,0056	.000,0639	.000,3702	.001,9451
15	.000,0003	.000,0034	.000,0396	.000,2389	.001,2990
16	.000,0002	.000,0021	.000,0251	.000,1574	.000,8822
17	.000,0001	.000,0013	.000,0162	.000,1056	.000,6085
18	.000,0001	.000,0008	.000,0107	.000,0721	.000,4259
19	.000,0000	.000,0005	.000,0071	.000,0500	.000,3020
20	.000,0000	.000,0004	.000,0049	.000,0351	.000,2169

$n = 11$

n' \ u'	17	18	19	20	21
11	.992,6683	.998,4095	.999,6853	.999,9688	.999,9972
12	.985,0134	.995,9965	.999,0474	.999,8610	.999,9830
13	.974,0303	.9918779	.997,8270	.999,5898	.999,9423
14	.959,7523	.985,7383	.995,8440	.999,0521	.999,8542
15	.942,4283	.977,4095	.992,9567	.998,1390	.999,6938
16	.922,4391	.966,8596	.989,0699	.996,7475	.999,4346
17	.900,2288	.954,1680	.984,1342	.994,7889	.999,0508
18	.876,2559	.939,4949	.978,1410	.992,1941	.998,5180
19	.850,9602	.923,0528	.971,1144	.988,9151	.997,8154
20	.824,7448	.905,0625	.963,1042	.984,9243	.996,9253

$n = 11$

n' \ u'	7	8	9	10	11
11	.022,6414	.063,4675	.134,9131	.259,9428	.409,9786
12	.014,9866	.044,2748	.099,1903	.201,6992	.334,9607
13	.010,1052	.031,2578	.073,5631	.156,8515	.273,4554
14	.006,9323	.022,3314	.055,0545	.122,4256	.223,4823
15	.004,8324	.016,1395	.041,5803	.095,9954	.183,0596
16	.003,4190	.011,7946	.031,6866	.075,6585	.150,4106
17	.002,4526	.008,7113	.024,3580	.059,9542	.124,0275
18	.001,7820	.006,4992	.018,8817	.047,7741	.102,6697
19	.001,3103	.004,8954	.014,7542	.038,2809	.085,3343
20	.000,9742	.003,7208	.011,6173	.030,8434	.071,2182

$n = 11$

n' \ u'	22
11	1.
12	.999,9993
13	.999,9952
14	.999,9825
15	.999,9529
16	.999,8953
17	.999,7966
18	.999,6423
19	.999,4174
20	.999,1074

TABLE I (*Continued*)

$n = 12$

$n-u'$	2	3	4	5	6
12	.000,0007	.000,0089	.000,0984	.000,5458	.002,7831
13	.000,0004	.000,0048	.000,0556	.000,3221	.001,7182
14	.000,0002	.000,0027	.000,0323	.000,1952	.001,0836
15	.000,0001	.000,0016	.000,0193	.000,1211	.000,6970
16	.000,0001	.000,0009	.000,0118	.000,0769	.000,4565
17	.000,0000	.000,0006	.000,0073	.000,0497	.000,3041
18	.000,0000	.000,0003	.000,0047	.000,0328	.000,2057
19	.000,0000	.000,0002	.000,0030	.000,0220	.000,1412
20	.000,0000	.000,0001	.000,0020	.000,0150	.000,0983

$n = 12$

$n-u'$	7	8	9	10	11
12	.009,4950	.029,6307	.069,9020	.150,4447	.263,2045
13	.006,1391	.020,0998	.049,7664	.112,5897	.206,8246
14	.004,0450	.013,8175	.035,8056	.084,6682	.162,8483
15	.002,7123	.009,6222	.026,0331	.064,0373	.128,6445
16	.001,8486	.006,7842	.019,1233	.048,7369	.102,0415
17	.001,2791	.004,8401	.014,1876	.037,3340	.081,3120
18	.000,8976	.003,4920	.010,6267	.028,7876	.065,1096
19	.000,6381	.002,5463	.008,0322	.022,3434	.052,3969
20	.000,4593	.001,8755	.006,1241	.017,4538	.042,3791

$n = 12$

$n-u'$	12	13	14	15	16
12	.421,0682	.578,9318	.736,7955	.849,5553	.930,0980
13	.347,5488	.500,0000	.664,1782	.793,1754	.893,6927
14	.285,9820	.429,6379	.593,8181	.734,5403	.851,8105
15	.235,0563	.368,0711	.527,6888	.675,9052	.806,2054
16	.193,2517	.314,8652	.466,8820	.618,8989	.758,5062
17	.159,0836	.269,2601	.411,8414	.564,6072	.710,0983
18	.131,2154	.230,3743	.362,5861	.513,6853	.662,0863
19	.108,4967	.197,3213	.318,8709	.466,4668	.615,3029
20	.089,9637	.169,2714	.280,3023	.423,0562	.570,3420

$n = 12$

$n-u'$	17	18	19	20	21
12	.970,3693	.990,5050	.997,2169	.999,4542	.999,9016
13	.950,2336	.981,6453	.993,8609	.998,5145	.999,6779
14	.925,1043	.969,0806	.988,6257	.996,7694	.999,2125
15	.895,7868	.952,7931	.981,2963	.993,9643	.998,3982
16	.863,2117	.933,0154	.971,7952	.989,8925	.997,1314
17	.828,3099	.910,1487	.960,1613	.984,4098	.995,3216
18	.791,9371	.884,6877	.946,5215	.977,4383	.992,8968
19	.754,8369	.857,1617	.931,0630	.968,9611	.989,8051
20	.717,6278	.828,0922	.914,0089	.959,0129	.986,0153

$n = 12$

$n-u'$	22	23	24
12	.999,9911	.999,9993	1.
13	.999,9571	.999,9952	.999,9998
14	.999,8640	.999,9825	.999,9987
15	.999,6650	.999,9529	.999,9948
16	.999,3030	.999,8953	.999,9850
17	.998,7164	.999,7966	.999,9649
18	.997,8435	.999,6423	.999,9285
19	.996,6268	.999,4174	.999,8685
20	.995,0161	.999,1074	.999,7768

$n = 13$

$n-u'$	2	3	4	5	6
13	.000,0002	.000,0025	.000,0302	.000,1825	.001,0201
14	.000,0001	.000,0013	.000,0169	.000,1063	.000,6196
15	.000,0001	.000,0007	.000,0097	.000,0636	.000,3844
16	.000,0000	.000,0004	.000,0057	.000,0389	.000,2431
17	.000,0000	.000,0003	.000,0035	.000,0243	.000,1566
18	.000,0000	.000,0002	.000,0021	.000,0155	.000,1025
19	.000,0000	.000,0001	.000,0013	.000,0100	.000,0682
20	.000,0000	.000,0001	.000,0009	.000,0066	.000,0460

TABLE I (*Continued*)

$n = 13$

n'' \ n'	7	8	9	10	11
13	.003,8123	.013,1194	.034,0605	.081,1780	.156,5660
14	.002,4162	.008,6899	.023,5900	.058,8796	.118,8720
15	.001,5607	.005,8382	.016,5321	.042,9993	.090,6403
16	.001,0260	.003,9760	.011,7198	.031,6825	.069,4665
17	.000,6856	.002,7431	.008,4011	.023,4462	.053,5364
18	.000,4652	.001,9158	.006,0864	.017,5103	.041,5003
19	.000,3201	.001,3537	.004,4545	.013,1753	.032,3612
20	.000,2232	.000,9671	.003,2917	.009,9865	.025,3845

$n = 13$

n'' \ n'	12	13	14	15	16
13	.277,1867	.417,9109	.582,0891	.722,8133	.843,4340
14	.220,5061	.347,5488	.505,6463	.652,4512	.787,9634
15	.175,3354	.288,2622	.436,4787	.584,6951	.729,8867
16	.139,5590	.238,8567	.375,1477	.521,1738	.671,3720
17	.111,3096	.197,9694	.321,5399	.462,7634	.614,0742
18	.089,0234	.164,2684	.275,1557	.409,8046	.559,1630
19	.071,4307	.136,5466	.235,3056	.362,2814	.507,3967
20	.057,5196	.113,7560	.201,2349	.319,9562	.459,2082

$n = 13$

n'' \ n'	17	18	19	20	21
13	.918,8220	.965,9395	.986,8806	.996,1877	.998,9799
14	.881,1280	.944,6493	.976,4100	.992,0943	.997,5838
15	.838,7805	.918,1821	.962,2942	.985,8206	.995,2312
16	.793,4081	.887,2820	.944,6493	.977,0996	.991,7022
17	.746,4711	.852,8615	.923,7885	.965,8193	.986,8347
18	.699,1865	.815,8728	.900,1462	.952,0067	.980,5300
19	.652,5120	.777,2204	.874,2159	.935,8003	.972,7510
20	.607,1635	.737,7122	.846,5029	.917,4183	.963,5133

$n = 13$

n'' \ n'	22	23	24	25	26
13	.999,8175	.999,9698	.999,9975	.999,9998	1.
14	.999,4659	.999,8937	.999,9870	.999,9987	.999,9999
15	.998,7601	.999,7226	.999,9559	.999,9948	.999,9996
16	.997,5432	.999,4017	.999,8845	.999,9850	.999,9985
17	.995,6611	.998,8707	.999,7461	.999,9649	.999,9953
18	.992,9766	.998,0683	.999,5084	.999,9285	.999,9885
19	.989,3787	.996,9308	.999,1355	.999,8685	.999,9753
20	.984,7879	.995,4252	.998,5900	.999,7768	.999,9527

$n = 14$

n'' \ n'	2	3	4	5	6
14	.000,0000	.000,0007	.000,0095	.000,0597	.000,3630
15	.000,0000	.000,0004	.000,0051	.000,0344	.000,2174
16	.000,0000	.000,0002	.000,0029	.000,0203	.000,1330
17	.000,0000	.000,0001	.000,0017	.000,0123	.000,0829
18	.000,0000	.000,0001	.000,0010	.000,0076	.000,0526
19	.000,0000	.000,0000	.000,0006	.000,0048	.000,0339
20	.000,0000	.000,0000	.000,0004	.000,0030	.000,0222

$n = 14$

n'' \ n'	7	8	9	10	11
14	.001,4751	.005,5531	.015,7478	.041,2348	.087,1113
15	.000,9191	.003,6036	.010,6505	.029,1065	.064,1731
16	.000,5835	.002,3732	.007,2948	.020,7174	.047,5626
17	.000,3770	.001,5849	.005,0577	.014,8721	.035,4823
18	.000,2476	.001,0726	.003,5478	.010,7670	.026,6493
19	.000,1651	.000,7351	.002,5165	.007,8606	.020,1520
20	.000,1116	.000,5098	.001,8039	.005,7858	.015,3423

TABLE I (*Continued*)

$n = 14$

$n-u'$	12	13	14	15	16
14	.169,6891	.279,7927	.426,5976	.573,4024	.720,2073
15	.130,6149	.224,7409	.357,6246	.500,0000	.651,8671
16	.100,7161	.180,4463	.298,5652	.433,5551	.585,4252
17	.077,8803	.145,0107	.248,6504	.374,4937	.522,5555
18	.060,4353	.116,7453	.206,8412	.322,6788	.464,2682
19	.047,0863	.094,2213	.172,0314	.277,6309	.411,0197
20	.036,8445	.076,2651	.143,1607	.238,7258	.363,9605

$n = 14$

$n-u'$	17	18	19	20	21	22
14	.830,3109	.912,8887	.958,7652	.984,2522	.994,4469	.998,5248
15	.775,2591	.874,9219	.935,8269	.972,7391	.989,3495	.996,7320
16	.715,3090	.832,2093	.908,1428	.957,3590	.981,9671	.993,7790
17	.661,3588	.786,2817	.876,5039	.938,1942	.972,1239	.989,3972
18	.605,8375	.738,5681	.841,8030	.915,5423	.959,7858	.983,3824
19	.552,7452	.690,3024	.804,9334	.889,8453	.945,0380	.975,6062
20	.502,7244	.642,4884	.766,7231	.861,6246	.928,0556	.966,0162

$n = 14$

$n-u'$	23	24	25	26	27	28
14	.999,6370	.999,9403	.999,9909	.999,9993	.999,9999	1.
15	.999,0809	.999,8131	.999,9656	.999,9961	.999,9996	.999,99999
16	.998,0742	.999,5385	.999,9046	.999,9859	.999,9985	.999,9999
17	.996,4635	.999,0331	.999,7826	.999,9610	.999,9953	.999,9995
18	.994,1081	.998,2033	.999,5684	.999,9097	.999,9885	.999,9986
19	.990,8904	.996,9535	.999,2272	.999,8166	.999,9753	.999,9963
20	.986,7220	.995,1925	.998,7219	.999,6631	.999,9527	.999,9916

$n = 15$

$n-u'$	2	3	4	5	6
15	.000,0000	.000,0002	.000,0027	.000,0131	.000,1259
16	.000,0000	.000,0001	.000,0015	.000,0109	.000,0745
17	.000,0000	.000,0001	.000,0008	.000,0064	.000,0450
18	.000,0000	.000,0000	.000,0005	.000,0038	.000,0277
19	.000,0000	.000,0000	.000,0003	.000,0023	.000,0173
20	.000,0000	.000,0000	.000,0002	.000,0014	.000,0110

$n = 15$

$n-u'$	7	8	9	10	11
15	.000,5530	.002,2613	.006,9592	.019,3785	.045,7170
16	.000,3395	.001,4416	.004,6103	.013,7030	.032,7977
17	.000,2123	.000,9329	.003,0948	.009,5355	.023,7050
18	.000,1351	.000,6124	.002,1040	.006,6980	.017,2643
19	.000,0873	.000,4074	.001,4477	.004,7484	.012,6703
20	.000,0573	.000,2745	.001,0076	.003,3967	.009,3695

$n = 15$

$n-u'$	12	13	14	15	16
15	.097,3340	.174,9095	.291,1827	.424,0664	.575,9336
16	.072,8057	.136,1517	.236,1717	.357,6246	.504,5928
17	.054,6203	.106,1457	.191,1627	.300,4703	.433,2736
18	.041,1534	.082,9593	.154,6265	.251,8892	.370,5976
19	.031,1546	.065,0425	.125,1165	.210,9365	.328,6325
20	.023,7043	.051,1792	.101,3509	.176,6085	.283,0953

TABLE I (*Continued*)

$n = 15$

n \ u'	17	18	19	20	21
15	.708,8173	.825,0905	.902,6060	.954,2830	.980,1215
16	.642,3754	.770,9725	.863,8483	.930,5283	.967,2023
17	.578,0768	.714,7113	.820,9525	.901,9511	.950,5323
18	.517,3504	.658,1253	.775,4377	.869,2877	.930,2901
19	.461,0404	.602,6434	.728,5128	.833,4039	.906,8276
20	.409,5455	.549,3124	.681,3118	.795,1935	.880,6049

$n = 15$

n \ u'	22	23	24	25	26
15	.993,0408	.997,7387	.999,4470	.999,8741	.999,9809
16	.987,2063	.995,3897	.998,6962	.999,6625	.999,9361
17	.978,8713	.991,7526	.997,3736	.999,2472	.999,8327
18	.967,8301	.986,6001	.995,2870	.998,5446	.999,6305
19	.954,0286	.979,7746	.992,2575	.997,4587	.999,2792
20	.937,5457	.971,1926	.988,1337	.995,8984	.998,7219

$n = 15$

n \ u'	27	28	29	30
15	.999,9973	.999,9998	.999,9999	1.
16	.999,9891	.999,9958	.999,9993	1.
17	.999,9679	.999,9856	.999,9935	1.
18	.999,9228	.999,9671	.999,9886	.999,9993
19	.999,8393	.999,9686	.999,9963	.999,9996
20	.999,6993	.999,9332	.999,9916	.999,9988

$n = 16$

n \ u'	2	3	4	5	6
16	.000,0000	.000,0001	.000,0008	.000,0060	.000,0427
17	.000,0000	.000,0000	.000,0004	.000,0034	.000,0250
18	.000,0000	.000,0000	.000,0002	.000,0020	.000,0149
19	.000,0000	.000,0000	.000,0001	.000,0012	.000,0091
20	.000,0000	.000,0000	.000,0001	.000,0007	.000,0056

$n = 16$

n \ u'	7	8	9	10	11
16	.000,2017	.000,8905	.002,9571	.009,1566	.022,7957
17	.000,1222	.000,5590	.001,9238	.006,1821	.015,9762
18	.000,0754	.000,3562	.001,2687	.004,2167	.011,2921
19	.000,0473	.000,2302	.000,8475	.002,9051	.008,0492
20	.000,0302	.000,1509	.000,5732	.002,0211	.005,7858

$n = 16$

n \ u'	12	13	14	15	16
16	.052,8017	.102,8117	.186,1617	.293,3260	.431,1087
17	.038,4600	.077,8067	.146,5073	.239,7439	.365,9287
18	.028,1549	.059,0702	.115,2798	.195,5792	.309,1455
19	.020,7242	.045,0178	.090,7885	.159,4445	.260,3267
20	.015,3423	.034,4554	.071,6196	.130,0205	.218,7595

TABLE I (*Continued*)

$n = 16$

n \ u'	17	18	19	20	21
16	.568,8913	.706,6740	.813,8383	.897,1883	.947,1983
17	.500,0000	.641,9579	.760,2561	.858,3998	.922,1933
18	.436,9076	.578,8655	.705,0503	.815,4620	.892,7501
19	.380,1245	.518,8376	.649,8444	.769,7200	.859,6266
20	.329,6833	.462,7919	.595,9004	.722,4357	.823,6640

$n = 16$

n \ u'	22	23	24	25	26
16	.977,2043	.990,8433	.997,0429	.999,1095	.999,7983
17	.963,4136	.984,0238	.994,2437	.998,0762	.999,4956
18	.945,7478	.974,6555	.989,9854	.996,3729	.998,9279
19	.924,3594	.962,6106	.984,0099	.993,8179	.997,9789
20	.899,5852	.947,8986	.976,1338	.990,2514	.996,5258

$n = 16$

n \ u'	27	28	29	30	31	32
16	.999,9573	.999,9940	.999,9992	.999,9999	1.	1.
17	.999,8778	.999,9786	.999,9966	.999,9997	1.	1.
18	.999,7140	.999,9408	.999,9894	.999,9986	.999,9999	1.
19	.999,4192	.999,8624	.999,9732	.999,9958	.999,9996	1.
20	.998,9391	.999,7188	.999,9415	.999,9893	.999,9988	.999,9999

$n = 17$

n \ u'	2	3	4	5	6
17	.000,0000	.000,0000	.000,0002	.000,0019	.000,0142
18	.000,0000	.000,0000	.000,0001	.000,0011	.000,0083
19	.000,0000	.000,0000	.000,0001	.000,0006	.000,0049
20	.000,0000	.000,0000	.000,0000	.000,0004	.000,0029

$n = 17$

n \ u'	7	8	9	10	11
17	.000,0718	.000,3406	.001,2141	.004,0529	.010,8662
18	.000,0430	.000,2109	.000,7773	.002,6855	.007,4596
19	.000,0262	.000,1325	.000,5046	.001,8001	.005,1685
20	.000,0163	.000,0845	.000,3318	.001,2189	.003,6139

$n = 17$

n \ u'	12	13	14	15	16
17	.027,2181	.057,1966	.112,1570	.190,6720	.302,8363
18	.019,3731	.042,2073	.085,8902	.151,4145	.249,4781
19	.013,8746	.031,3856	.065,8689	.120,2125	.204,9038
20	.010,0005	.023,3061	.050,6268	.095,5108	.167,9944

$n = 17$

n \ u'	17	18	19	20	21
17	.429,0211	.570,9789	.697,1637	.809,3280	.887,8430
18	.365,9287	.503,8306	.634,0713	.756,6508	.848,5855
19	.310,7679	.441,7747	.572,7816	.702,1711	.805,6826
20	.263,1290	.385,4450	.514,5563	.647,4428	.760,3963

TABLE I (*Continued*)

$m = 17$

n \ u'	22	23	24	25	26
17	.942,8034	.972,7819	.989,1338	.995,9471	.998,7859
18	.917,2300	.957,7927	.981,6197	.992,5404	.997,5044
19	.887,1980	.939,0714	.971,4081	.987,5764	.995,4361
20	.853,4169	.916,8400	.958,3533	.980,8396	.992,3711

$m = 17$

n \ u'	27	28	29	30	31
17	.999,6594	.999,9282	.999,9858	.999,9981	.999,9998
18	.999,2227	.999,8101	.999,9570	.999,9929	.999,9989
19	.998,4590	.999,5751	.999,8940	.999,9795	.999,9965
20	.997,2498	.999,1604	.999,7745	.999,9499	.999,9909

$m = 17$

n \ u'	32	33	34
17	1.	1.	1.
18	.999,9999	1.	1.
19	.999,9996	1.	1.
20	.999,9986	.999,9999	1.

$m = 18$

n \ u'	2	3	4	5	6
18	.000,0000	.000,0000	.000,0001	.000,0006	.000,0047
19	.000,0000	.000,0000	.000,0000	.000,0003	.000,0027
20	.000,0000	.000,0000	.000,0000	.000,0002	.000,0016

$m = 18$

n \ u'	7	8	9	10	11
18	.000,0250	.000,1269	.000,4836	.001,7319	.004,9776
19	.000,0148	.000,0776	.000,3053	.001,1295	.003,3548
20	.000,0090	.000,0482	.000,1954	.000,7448	.002,2533

$m = 18$

n \ u'	12	13	14	15	16
18	.013,4163	.030,2938	.064,0488	.117,0923	.200,4463
19	.009,3549	.021,8551	.047,8555	.090,5705	.160,6125
20	.006,5691	.015,8550	.035,8553	.070,1416	.128,5099

$m = 18$

n \ u'	17	18	19	20	21
18	.304,6389	.434,8796	.565,1204	.695,3611	.799,5537
19	.252,5426	.372,9273	.500,0000	.633,7607	.747,4574
20	.208,7664	.318,2070	.439,8077	.573,5684	.693,9531

$m = 18$

n \ u'	22	23	24	25	26
18	.882,9077	.935,9512	.969,7062	.986,5837	.995,0224
19	.843,7651	.909,4295	.954,0017	.978,1449	.991,1451
20	.800,9617	.878,7861	.934,5013	.967,0019	.985,5736

$m = 18$

n \ u'	27	28	29	30	31
18	.998,2681	.999,5164	.999,8731	.999,9750	.999,9954
19	.996,6452	.998,9530	.999,6947	.999,9302	.999,9852
20	.994,1452	.997,9914	.999,3651	.999,8360	.999,9616

$m = 18$

n \ u'	32	33	34	35	36
18	.999,9994	.999,9999	1.	1.	1.
19	.999,9977	.999,9997	1.	1.	1.
20	.999,9930	.999,9989	.999,9999	.999,99999	1.

TABLE I (*Concluded*)

$n = 19$

n \ u'	2	3	4	5	6
19	.000,0000	.000,0000	.000,0000	.000,0002	.000,0015
20	.000,0000	.000,0000	.000,0000	.000,0001	.000,0009

$n = 19$

n \ u'	7	8	9	10	11
19	.000,0086	.000,0462	.000,1875	.000,7174	.002,2009
20	.000,0050	.000,0280	.000,1169	.000,4611	.001,4591

$n = 19$

n \ u'	12	13	14	15	16
19	.006,3548	.015,3550	.034,8553	.068,2844	.125,5915
20	.004,3501	.010,8549	.025,4705	.051,5699	.098,1013

$n = 19$

n \ u'	17	18	19	20	21
19	.204,3888	.312,7350	.433,1196	.566,8804	.687,2650
20	.164,9901	.260,9611	.372,9273	.503,2583	.627,0727

$n = 19$

n \ u'	22	23	24	25	26
19	.795,6112	.874,4085	.931,7156	.965,1447	.984,6450
20	.744,3706	.835,0099	.904,8070	.948,4301	.975,5734

$n = 19$

n \ u'	27	28	29	30	31
19	.993,5452	.997,7991	.999,2825	.999,8125	.999,9538
20	.989,1451	.995,8908	.998,5409	.999,5734	.999,8331

$n = 19$

n \ u'	32	33	34	35	36
19	.999,9914	.999,9985	.999,9998	.999,9999	1.
20	.999,9749	.999,9950	.999,9993	.999,9999	1.

$n = 20$

n \ u'	2	3	4	5	6
20	.000,0000	.000,0000	.000,0000	.000,0001	.000,0

$n = 20$

n \ u'	7	8	9	10	11
20	.000,0029	.000,0165	.000,0710	.000,2890	.000,1

$n = 20$

n \ u'	12	13	14	15	16
20	.002,9046	.007,4821	.018,1627	.037,9982	.07

$n = 20$

n \ u'	17	18	19	20	21
20	.130,0916	.212,9756	.314,2784	.438,0928	.56

$n = 20$

n \ u'	22	23	24	25	26
20	.685,7216	.787,0244	.869,9084	.925,1644	.96

$n = 20$

n \ u'	27	28	29	30	31
20	.981,8373	.992,5179	.997,0954	.999,0571	.99

$n = 20$

n \ u'	32	33	34	35	36
20	.999,9290	.999,9835	.999,9971	.999,9995	.99

$n = 20$

n \ u'	37	38	39	40
20	1.	1.	1.	1.

TABLE II

TABLE II

Significance Levels of u

When $\varepsilon < .50$, u_ε is the largest integer, u', for which $P\{u \leq u'\} \leq \varepsilon$; when $\varepsilon > .50$, u is the smallest integer, u', for which $P\{u \leq u'\} \geq \varepsilon$.

$m = 2$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
2					4	4	4	4
3					5	5	5	5
4					5	5	5	5
5					5	5	5	5
6					5	5	5	5
7					5	5	5	5
8				2	5	5	5	5
9				2	5	5	5	5
10				2	5	5	5	5
11				2	5	5	5	5
12			2	2	5	5	5	5
13			2	2	5	5	5	5
14			2	2	5	5	5	5
15			2	2	5	5	5	5
16			2	2	5	5	5	5
17			2	2	5	5	5	5
18			2	2	5	5	5	5
19		2	2	2	5	5	5	5
20		2	2	2	5	5	5	5

$m = 3$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
3					6	6	6	6
4					6	7	7	7
5				2	7	7	7	7
6			2	2	7	7	7	7
7			2	2	7	7	7	7
8			2	2	7	7	7	7
9		2	2	2	7	7	7	7
10		2	2	3	7	7	7	7
11		2	2	3	7	7	7	7
12	2	2	2	3	7	7	7	7
13	2	2	2	3	7	7	7	7
14	2	2	2	3	7	7	7	7
15	2	2	3	3	7	7	7	7
16	2	2	3	3	7	7	7	7
17	2	2	3	3	7	7	7	7
18	2	2	3	3	7	7	7	7
19	2	2	3	3	7	7	7	7
20	2	2	3	3	7	7	7	7

$m = 4$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
4				2	7	8	8	8
5			2	2	8	8	8	9
6		2	2	3	8	8	9	9
7		2	2	3	8	9	9	9
8	2	2	3	3	9	9	9	9
9	2	2	3	3	9	9	9	9
10	2	2	3	3	9	9	9	9
11	2	2	3	3	9	9	9	9
12	2	3	3	4	9	9	9	9
13	2	3	3	4	9	9	9	9
14	2	3	3	4	9	9	9	9
15	3	3	3	4	9	9	9	9
16	3	3	4	4	9	9	9	9
17	3	3	4	4	9	9	9	9
18	3	3	4	4	9	9	9	9
19	3	3	4	4	9	9	9	9
20	3	3	4	4	9	9	9	9

$m = 5$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
5		2	2	3	8	9	9	10
6	2	2	3	3	9	9	10	10
7	2	2	3	3	9	10	10	11
8	2	2	3	3	10	10	11	11
9	2	3	3	4	10	11	11	11
10	3	3	3	4	10	11	11	11
11	3	3	4	4	11	11	11	11
12	3	3	4	4	11	11	11	11
13	3	3	4	4	11	11	11	11
14	3	3	4	5	11	11	11	11
15	3	4	4	5	11	11	11	11
16	3	4	4	5	11	11	11	11
17	3	4	4	5	11	11	11	11
18	4	4	5	5	11	11	11	11
19	4	4	5	5	11	11	11	11
20	4	4	5	5	11	11	11	11

TABLE I (*Concluded*)

m = 19

n \ u'	2	3	4	5	6
19	.000,0000	.000,0000	.000,0000	.000,0002	.000,0015
20	.000,0000	.000,0000	.000,0000	.000,0001	.000,0009

m = 19

n \ u'	7	8	9	10	11
19	.000,0086	.000,0462	.000,1875	.000,7174	.002,2009
20	.000,0050	.000,0280	.000,1169	.000,4611	.001,4591

m = 19

n \ u'	12	13	14	15	16
19	.006,3548	.015,3550	.034,8553	.068,2844	.125,5915
20	.004,3501	.010,8549	.025,4705	.051,5699	.098,1013

m = 19

n \ u'	17	18	19	20	21
19	.204,3888	.312,7350	.433,1196	.566,8804	.687,2650
20	.164,9901	.260,9611	.372,9273	.503,2583	.627,0727

m = 19

n \ u'	22	23	24	25	26
19	.795,6112	.874,4085	.931,7156	.965,1447	.984,6450
20	.744,3706	.835,0093	.904,8070	.948,4301	.975,5734

m = 19

n \ u'	27	28	29	30	31
19	.993,6452	.997,7991	.999,2825	.999,8125	.999,9538
20	.989,1451	.995,8908	.998,5409	.999,5734	.999,8831

m = 19

n \ u'	32	33	34	35	36
19	.999,9914	.999,9985	.999,9998	.999,9999	1.
20	.999,9749	.999,9950	.999,9993	.999,9999	1.

m = 20

n \ u'	2	3	4	5	6
20	.000,0000	.000,0000	.000,0000	.000,0001	.000,0005

m = 20

n \ u'	7	8	9	10	11
20	.000,0029	.000,0165	.000,0710	.000,2890	.000,9429

m = 20

n \ u'	12	13	14	15	16
20	.002,9046	.007,4821	.018,1627	.037,9982	.074,8356

m = 20

n \ u'	17	18	19	20	21
20	.130,0916	.212,9756	.314,2784	.438,0928	.561,9072

m = 20

n \ u'	22	23	24	25	26
20	.685,7216	.787,0244	.869,9084	.925,1644	.962,0018

m = 20

n \ u'	27	28	29	30	31
20	.981,8373	.992,5179	.997,0954	.999,0571	.999,7110

m = 20

n \ u'	32	33	34	35	36
20	.999,9290	.999,9835	.999,9971	.999,9995	.999,9999

m = 20

n \ u'	37	38	39	40
20	1.	1.	1.	1.

TABLE II

TABLE II

Significance Levels of u

When $\varepsilon < .50$, u_ε is the largest integer, u', for which $P\{u \leq u'\} \leq \varepsilon$; when $\varepsilon > .50$, u is the smallest integer, u', for which $P\{u \leq u'\} \geq \varepsilon$.

m = 2

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
2					4	4	4	4
3					5	5	5	5
4					5	5	5	5
5					5	5	5	5
6					5	5	5	5
7					5	5	5	5
8				2	5	5	5	5
9				2	5	5	5	5
10				2	5	5	5	5
11				2	5	5	5	5
12			2	2	5	5	5	5
13			2	2	5	5	5	5
14			2	2	5	5	5	5
15			2	2	5	5	5	5
16			2	2	5	5	5	5
17			2	2	5	5	5	5
18			2	2	5	5	5	5
19		2	2	2	5	5	5	5
20		2	2	2	5	5	5	5

m = 3

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
3					6	6	6	6
4					6	7	7	7
5				2	7	7	7	7
6			2	2	7	7	7	7
7			2	2	7	7	7	7
8			2	2	7	7	7	7
9		2	2	2	7	7	7	7
10		2	2	3	7	7	7	7
11		2	2	3	7	7	7	7
12	2	2	2	3	7	7	7	7
13	2	2	2	3	7	7	7	7
14	2	2	2	3	7	7	7	7
15	2	2	3	3	7	7	7	7
16	2	2	3	3	7	7	7	7
17	2	2	3	3	7	7	7	7
18	2	2	3	3	7	7	7	7
19	2	2	3	3	7	7	7	7
20	2	2	3	3	7	7	7	7

m = 4

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
4				2	7	8	8	8
5			2	2	8	8	8	9
6		2	2	3	8	8	9	9
7		2	2	3	8	9	9	9
8	2	2	3	3	9	9	9	9
9	2	2	3	3	9	9	9	9
10	2	2	3	3	9	9	9	9
11	2	2	3	3	9	9	9	9
12	2	3	3	4	9	9	9	9
13	2	3	3	4	9	9	9	9
14	2	3	3	4	9	9	9	9
15	3	3	3	4	9	9	9	9
16	3	3	4	4	9	9	9	9
17	3	3	4	4	9	9	9	9
18	3	3	4	4	9	9	9	9
19	3	3	4	4	9	9	9	9
20	3	3	4	4	9	9	9	9

m = 5

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
5		2	2	3	8	9	9	10
6	2	2	3	3	9	9	10	10
7	2	2	3	3	9	10	10	11
8	2	2	3	3	10	10	11	11
9	2	3	3	4	10	11	11	11
10	3	3	3	4	10	11	11	11
11	3	3	4	4	11	11	11	11
12	3	3	4	4	11	11	11	11
13	3	3	4	4	11	11	11	11
14	3	3	4	5	11	11	11	11
15	3	4	4	5	11	11	11	11
16	3	4	4	5	11	11	11	11
17	3	4	4	5	11	11	11	11
18	4	4	5	5	11	11	11	11
19	4	4	5	5	11	11	11	11
20	4	4	5	5	11	11	11	11

TABLE II (*Continued*)

$m = 6$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
6	2	2	3	3	10	10	11	11
7	2	3	3	4	10	11	11	12
8	3	3	3	4	11	11	12	12
9	3	3	4	4	11	12	12	13
10	3	3	4	5	11	12	13	13
11	3	4	4	5	12	12	13	13
12	3	4	4	5	12	12	13	13
13	3	4	5	5	12	13	13	13
14	4	4	5	5	12	13	13	13
15	4	4	5	6	13	13	13	13
16	4	4	5	6	13	13	13	13
17	4	5	5	6	13	13	13	13
18	4	5	5	6	13	13	13	13
19	4	5	6	6	13	13	13	13
20	4	5	6	6	13	13	13	13

$m = 7$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
7	3	3	3	4	11	12	12	12
8	3	3	4	4	12	12	13	13
9	3	4	4	5	12	13	13	14
10	3	4	5	5	12	13	14	14
11	4	4	5	5	13	13	14	14
12	4	4	5	6	13	13	14	15
13	4	5	5	6	13	14	15	15
14	4	5	5	6	13	14	15	15
15	4	5	6	6	14	14	15	15
16	5	5	6	6	14	15	15	15
17	5	5	6	7	14	15	15	15
18	5	5	6	7	14	15	15	15
19	5	6	6	7	14	15	15	15
20	5	6	6	7	14	15	15	15

$m = 8$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
8	3	4	4	5	12	13	13	14
9	3	4	5	5	13	13	14	14
10	4	4	5	6	13	14	14	15
11	4	5	5	6	14	14	15	15
12	4	5	6	6	14	15	15	16
13	5	5	6	6	14	15	16	16
14	5	5	6	7	15	15	16	16
15	5	5	6	7	15	15	16	17
16	5	6	6	7	15	16	16	17
17	5	6	7	7	15	16	17	17
18	6	6	7	8	15	16	17	17
19	6	6	7	8	15	16	17	17
20	6	6	7	8	16	16	17	17

$m = 9$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
9	4	4	5	6	13	14	15	15
10	4	5	5	6	14	15	15	16
11	5	5	6	6	14	15	16	16
12	5	5	6	7	15	15	16	17
13	5	6	6	7	15	16	17	17
14	5	6	7	7	16	16	17	17
15	6	6	7	8	16	17	17	18
16	6	6	7	8	16	17	17	18
17	6	7	7	8	16	17	18	18
18	6	7	8	8	17	17	18	19
19	6	7	8	8	17	17	18	19
20	7	7	8	9	17	17	18	19

TABLE II (*Continued*)

$m = 10$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
10	5	5	6	6	15	15	16	16
11	5	5	6	7	15	16	17	17
12	5	6	7	7	16	16	17	18
13	5	6	7	8	16	17	18	18
14	6	6	7	8	16	17	18	18
15	6	7	7	8	17	17	18	19
16	6	7	8	8	17	18	19	19
17	7	7	8	9	17	18	19	19
18	7	7	8	9	18	18	19	20
19	7	8	8	9	18	19	19	20
20	7	8	9	9	18	19	19	20

$m = 11$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
11	5	6	7	7	16	16	17	18
12	6	6	7	8	16	17	18	18
13	6	6	7	8	17	18	18	19
14	6	7	8	8	17	18	19	19
15	7	7	8	9	18	18	19	20
16	7	7	8	9	18	19	20	20
17	7	8	9	9	18	19	20	21
18	7	8	9	10	19	19	20	21
19	8	8	9	10	19	20	21	21
20	8	8	9	10	19	20	21	21

$m = 12$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
12	6	7	7	8	17	18	18	19
13	6	7	8	9	17	18	19	20
14	7	7	8	9	18	19	20	20
15	7	8	8	9	18	19	20	21
16	7	8	9	10	19	20	21	21
17	8	8	9	10	19	20	21	21
18	8	8	9	10	20	20	21	22
19	8	9	10	10	20	21	22	22
20	8	9	10	11	20	21	22	22

$m = 13$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
13	7	7	8	9	18	19	20	20
14	7	8	9	9	19	19	20	21
15	7	8	9	10	19	20	21	21
16	8	8	9	10	20	20	21	22
17	8	9	10	10	20	21	22	22
18	8	9	10	11	20	21	22	23
19	9	9	10	11	21	22	23	23
20	9	10	10	11	21	22	23	23

TABLE II (*Concluded*)

$n = 14$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
14	7	8	9	10	19	20	21	22
15	8	8	9	10	20	21	22	22
16	8	9	10	11	20	21	22	23
17	8	9	10	11	21	22	23	23
18	9	9	10	11	21	22	23	24
19	9	10	11	12	22	22	23	24
20	9	10	11	12	22	23	24	24

$n = 15$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
15	8	9	10	11	20	21	22	23
16	9	9	10	11	21	22	23	23
17	9	10	11	11	21	22	23	24
18	9	10	11	12	22	23	24	24
19	10	10	11	12	22	23	24	25
20	10	11	12	12	23	24	25	25

$n = 16$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
16	9	10	11	11	22	22	23	24
17	9	10	11	12	22	23	24	25
18	10	10	11	12	23	24	25	25
19	10	11	12	13	23	24	25	26
20	10	11	12	13	24	24	25	26

$n = 17$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
17	10	10	11	12	23	24	25	25
18	10	11	12	13	23	24	25	26
19	10	11	12	13	24	25	26	26
20	11	11	13	13	24	25	26	27

$n = 18$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
18	11	11	12	13	24	25	26	26
19	11	12	13	14	24	25	26	27
20	11	12	13	14	25	26	27	28

$n = 19$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
19	11	12	13	14	25	26	27	28
20	12	12	13	14	26	26	28	28

$n = 20$

n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
19								
20	12	13	14	15	26	27	28	29

TABLE III

Significance Levels of u

$$P\{u \leq u_\varepsilon\} = \varepsilon$$

For definition of u_ε, see note to Table II.

m=n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
10	4	5	6	6	15	15	16	17
11	5	6	7	7	16	16	17	18
12	6	6	7	8	17	18	19	19
13	7	7	8	9	18	19	20	20
14	7	8	9	10	19	20	21	22
15	8	9	10	11	20	21	22	23
16	9	10	11	11	22	22	23	24
17	10	10	11	12	23	24	25	25
18	10	11	12	13	24	25	26	27
19	11	12	13	14	25	26	27	28
20	12	13	14	15	26	27	28	29
21	13	14	15	16	27	28	29	30
22	14	14	16	17	28	29	31	31
23	14	15	16	17	30	31	32	33
24	15	16	17	18	31	32	33	34
25	16	17	18	19	32	33	34	35
26	17	18	19	20	33	34	35	36
27	18	19	20	21	34	35	36	37
28	18	19	21	22	35	36	38	39
29	19	20	22	23	36	37	39	40
30	20	21	22	24	37	39	40	41
31	21	22	23	25	38	40	41	42
32	22	23	24	25	40	41	42	43
33	23	24	25	26	41	42	43	44
34	23	24	26	27	42	43	45	46
35	24	25	27	28	43	44	46	47
36	25	26	28	29	44	45	47	48
37	26	27	29	30	45	46	48	49
38	27	28	30	31	46	47	49	50
39	28	29	30	32	47	49	50	51
40	29	30	31	33	48	50	51	52
41	29	31	32	34	49	51	52	54
42	30	31	33	35	50	52	54	55
43	31	32	34	35	52	53	55	56
44	32	33	35	36	53	54	56	57
45	33	34	36	37	54	55	57	58
46	34	35	37	38	55	56	58	59
47	35	36	38	39	56	57	59	60
48	35	37	38	40	57	59	60	62
49	36	38	39	41	58	60	61	63
50	37	38	40	42	59	61	63	64
51	38	39	41	43	60	62	64	65
52	39	40	42	44	61	63	65	66
53	40	41	43	45	62	64	66	67
54	41	42	44	45	64	65	67	68
55	42	43	45	46	65	66	68	69

m=n	$u_{.005}$	$u_{.01}$	$u_{.025}$	$u_{.05}$	$u_{.95}$	$u_{.975}$	$u_{.99}$	$u_{.995}$
56	42	44	46	47	66	67	69	71
57	43	45	47	48	67	68	70	72
58	44	46	47	49	68	70	71	73
59	45	46	48	50	69	71	73	74
60	46	47	49	51	70	72	74	75
61	47	48	50	52	71	73	75	76
62	48	49	51	53	72	74	76	77
63	49	50	52	54	73	75	77	78
64	49	51	53	55	74	76	78	80
65	50	52	54	56	75	77	79	81
66	51	53	55	57	76	78	80	82
67	52	54	56	58	77	79	81	83
68	53	54	57	58	79	80	83	84
69	54	55	58	59	80	81	84	85
70	55	56	58	60	81	83	85	86
71	56	57	59	61	82	84	86	87
72	57	58	60	62	83	85	87	88
73	57	59	61	63	84	86	88	90
74	58	60	62	64	85	87	89	91
75	59	61	63	65	86	88	90	92
76	60	62	64	66	87	89	91	93
77	61	63	65	67	88	90	92	94
78	62	64	66	68	89	91	93	95
79	63	64	67	69	90	92	95	96
80	64	65	68	70	91	93	96	97
81	65	66	69	71	92	94	97	98
82	66	67	69	71	94	96	98	99
83	66	68	70	72	95	97	99	101
84	67	69	71	73	96	98	100	102
85	68	70	72	74	97	99	101	103
86	69	71	73	75	98	100	102	104
87	70	72	74	76	99	101	103	105
88	71	73	75	77	100	102	104	106
89	72	74	76	78	101	103	105	107
90	73	74	77	79	102	104	107	108
91	74	75	78	80	103	105	108	109
92	75	76	79	81	104	106	109	110
93	75	77	80	82	105	107	110	112
94	76	78	81	83	106	108	111	113
95	77	79	82	84	107	109	112	114
96	78	80	82	85	108	111	113	115
97	79	81	83	86	109	112	114	116
98	80	82	84	87	110	113	115	117
99	81	83	85	87	112	114	116	118
100	82	84	86	88	113	115	117	119

NOTES

This section is devoted to brief research and expository articles, and notes on methodology.

A NOTE ON THE BEST LINEAR ESTIMATE

By Allen T. Craig

University of Iowa

1. Introduction. Let the chance variable x be subject to the distribution function $D(x)$ and as usual let $E[g(x)]$ denote the mathematical expectation of the function $g(x)$. If $x_1, x_2, \cdots, x_n$ constitute a sample of n independent values of x, the function $y = c_1x_1 + c_2x_2 + \cdots + c_nx_n$ is frequently called the best linear estimate of $E(x)$ when the c's are so chosen that $E(y) = E(x)$, and $E[y - E(x)]^2 = \sigma_y^2$ is a minimum. It is the purpose of this note to give an example of an estimate y, best in the sense defined, yet such that, y' being another estimate,

$$Pr[E(x) - \delta \leq y \leq E(x) + \delta] \leq Pr[E(x) - \delta \leq y' \leq E(x) + \delta],$$

for every $\delta > 0$.

2. The rectangular distribution. Consider $D(x) = 1/a$, $0 \leq x \leq a$, and let the n items of each sample be arranged in ascending order of magnitude so that $x_1 \leq x_2 \leq \cdots \leq x_n$, $n \geq 2$. The generating function $G(t)$ of the moments of the distribution of $y = c_1x_1 + c_2x_2 + \cdots + c_nx_n$ is

$$G(t) = E(e^{ty}) = \frac{n!}{a^n} \int_0^a \int_0^{x_n} \int_0^{x_{n-1}} \cdots \int_0^{x_2} e^{t(c_1x_1 + \cdots + c_nx_n)}\, dx_1\, dx_2 \cdots dx_n .$$

Thus

$$E(y) = G'(0) = \frac{a}{n+1}[c_1 + 2c_2 + 3c_3 + \cdots + nc_n],$$

and

$$\begin{aligned} E(y^2) = G''(0) = \frac{a^2}{(n+1)(n+2)} [1 \cdot 2c_1^2 + 2 \cdot 3c_2^2 + \cdots + n(n+1)c_n^2 \\ + 2\{1 \cdot 3c_1c_2 + 1 \cdot 4c_1c_3 + \cdots + 1 \cdot (n+1)c_1c_n \\ + 2 \cdot 4c_2c_3 + \cdots + 2(n+1)c_2c_n \\ \vdots \\ + (n-1)(n+1)c_{n-1}c_n\}]. \end{aligned}$$

From $E(y) = E(x) = a/2$, we have

$$c_1 = \tfrac{1}{2}(n+1) - 2c_2 - \cdots - nc_n .$$

Thus $\sigma_y^2 = G''(0) - a^2/4$ with c_1 in $G''(0)$ replaced by $\frac{1}{2}(n+1) - 2c_2 - \cdots - nc_n$. From $\dfrac{\partial \sigma_y^2}{\partial c_j} = 0$, $j = 2, 3, \cdots, n$, we obtain the following system of $n-1$ non-homogeneous linear equations in $n-1$ unknowns:

$$\begin{aligned}
4c_2 + 6c_3 + \cdots + 2nc_n &= n+1 \\
6c_2 + 12c_3 + \cdots + 4nc_n &= 2(n+1) \\
8c_2 + 16c_3 + \cdots + 6nc_n &= 3(n+1) \\
&\vdots \\
2nc_2 + 4nc_3 + \cdots + 2n(n-1)c_n &= (n-1)(n+1).
\end{aligned}$$

Since the determinant of the coefficients is not zero, the solution $c_2 = c_3 = \cdots = c_{n-1} = 0$, $c_n = (n+1)/2n$, is unique. Further, we see that $c_1 = 0$ so the best linear estimate of the mean of the rectangular population is $y = (n+1)x_n/2n$, where x_n is the largest item in the sample.

The distribution function of y is readily found to be

$$D(y) = n\left[\frac{2n}{a(n+1)}\right]^n y^{n-1}, \qquad 0 \le y \le \frac{n+1}{2n}\,a.$$

From this, it follows that $\sigma_y^2 = \dfrac{a^2}{4n(n+2)}$.

It has long been known[1] that the sampling distribution of the statistic $\omega = \frac{1}{2}(x_1 + x_n)$, where x_1 and x_n are respectively the smallest and largest items in samples of size n from a rectangular population, has a smaller variance than does that of the arithmetic mean $\bar{x}$ of all n items. The distribution function of ω is

$$\begin{aligned}
D(\omega) &= \frac{2^{n-1} n \omega^{n-1}}{a^n}, && 0 \le \omega \le \tfrac{1}{2}a, \\
&= \frac{2^{n-1} n}{a^n}(a - \omega)^{n-1}, && \tfrac{1}{2}a \le \omega < a,
\end{aligned}$$

so that $E(\omega) = \frac{1}{2}a$ and $\sigma_\omega^2 = \dfrac{a^2}{2(n+1)(n+2)}$. Thus $\sigma_y^2 = \frac{1}{2}\sigma_\omega^2$, approximately. Yet Pittman has recently proved that for every $\delta > 0$, $Pr[E(x) - \delta \le \omega \le E(x) + \delta]$ exceeds the probability that any other estimate, including y, will fall in this interval of length 2δ about the mean $a/2$.

If we write $u = \dfrac{y - a/2}{\sigma_y}$ and $v = \dfrac{\omega - a/2}{\sigma_\omega}$, then the limits of $D(u)$ and $D(v)$ as n approaches infinity are respectively e^{u-1}, $-\infty \le u \le 1$, and $\dfrac{1}{\sqrt{2}}\, e^{-\sqrt{2}|v|}$,

[1] R. A. Fisher, "Theoretical foundations of mathematical statistics," *Phil. Trans. Roy. Soc. London*, Series A, Vol. 222 (1921), pp. 309–368.

$-\infty \leq v \leq \infty$. Thus neither y nor ω has an asymptotic normal distribution. It is, of course, this fact which makes the criterion of minimum variance illusory.

3. Other polynomial distribution functions. Let repeated samples of n independent values of x be drawn from a population characterized by $D(x) = \frac{k+1}{a^{k+1}} x^k$, $0 \leq x \leq a$, and k a positive integer or zero. It can be shown that the best linear estimate of the mean of the population is $y = \frac{(k+1)n+1}{n(k+2)} x_n$, where as before x_n is the largest item of the sample. The sampling distribution of y is easily obtained. It follows that

$$\sigma_y^2 = \frac{(k+1)a^2}{(k+2)^2[(k+1)n^2+2n]} = \frac{k+3}{n(k+1)+2}\sigma_{\bar{x}}^2,$$

where as usual $\bar{x}$ is the arithmetic mean of the sample. Again, if we write $u = \left(y - \frac{k+1}{k+2}a\right)\Big/\sigma_y$, the limit of the distribution of u as n approaches infinity is, as before, e^{u-1}, $-\infty \leq u \leq 1$.

A NOTE ON TOLERANCE LIMITS

By Edward Paulson[1]

Columbia University

Among various statistical problems arising in the process of controlling quality in mass production, a rather important one appears to be the determination of tolerance limits when the variability of the product is known to be due to random factors. This problem was recently treated in a pioneer article by Wilks. This note will point out a relationship between tolerance limits and confidence limits (used in the sense of Neyman), and will use this concept to establish tolerance limits when the product is described by two qualities, the measurements on which are assumed to have a bivariate normal distribution.

For the case of a single variate, the problem of finding tolerance limits as stated by Wilks is to find a sample size n, and two functions $L_1(x_1 \cdots x_n)$ and $L_2(x_1 x_2 \cdots x_n)$ so that if $P = \int_{L_1}^{L_2} f(x)\,dx$ denotes the conditional probability of a future observation falling between the random variates L_2 and L_1, then

$$E(P) = \alpha, \quad \text{and Prob.} \quad [\alpha - \Delta_1 \leq P \leq \alpha + \Delta_2] \geq \beta.$$

The relationship between confidence limits and tolerance limits will arise if confidence limits are determined, not for a parameter of the distribution, but for

---

[1] Work done under a grant-in-aid from the Carnegie Corporation of New York.

a future random observation (or for some function of the observations in a future independent sample). This is based on the following simple lemma: *If confidence limits* $U_1(x_1 \cdots x_n)$ *and* $U_2(x_1 \cdots x_n)$ *on a probability level* $= \alpha_0$ *are determined for* g, *a function of a future sample of* k *observations, and* $P = \int_{U_1}^{U_2} \psi(g)\, dg$, *then* $E(P) = \alpha_0$. For let $\psi(g)\, dg$ and $\varphi(U_1, U_2)\, dU_1\, dU_2$ denote the distribution of g and U_1, U_2 respectively, then by the definition of expected value

$$E(P) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \left[\int_{U_1}^{U_2} \psi(g)\, dg \right] \varphi(U_1, U_2)\, dU_1\, dU_2 .$$

This triple integral is however exactly the probability that g will lie between U_1 and U_2, which by the nature of confidence limits must equal α_0, which proves the lemma. In a similar manner it follows that if on the basis of a given sample an l dimensional confidence region is found for statistics $g_1, g_2, \cdots g_l$ derived from a future sample, and if P denotes the probability that $g_1 \cdots g_l$ all fall in the confidence region, then $E(P)$ in repeated sampling equals α. To establish tolerance limits, it is necessary in addition to $E(P)$ to also know the distribution of P, or at least σ_P^2, so the distribution of P can be approximated.

It appears, at least on an intuitive basis, that the "best" confidence interval can be used to determine the shape of the "most efficient" tolerance limits; this intuitive notion will gain additional support from the character of the tolerance region which will now be derived for an observation (x, y) from a distribution with probability density $f(x, y)$, where

$$f(x, y) = \frac{\exp\left\{ -\frac{1}{2(1-\rho^2)} \left[\left(\frac{x - m_x}{\sigma_x} \right)^2 - 2\rho \left(\frac{x - m_x}{\sigma_x} \right) \left(\frac{y - m_y}{\sigma_y} \right) + \left(\frac{y - m_y}{\sigma_y} \right)^2 \right] \right\}}{2\pi \sigma_x \sigma_y \sqrt{1 - \rho^2}}$$

Suppose we have 2 independent samples

$$[(x_1, y_1)(x_2, y_2) \cdots (x_n, y_n)] \quad \text{and} \quad [(x, y)]$$

both from $f(x, y)$. Then it is known that

$$T^2 = \left(\frac{n}{n+1} \right) \frac{1}{1 - r^2} \left\{ \left(\frac{\bar{x} - x}{s_x} \right)^2 - \frac{2r}{s_x s_y} (\bar{x} - x)(\bar{y} - y) + \frac{(\bar{y} - y)^2}{s^2 y} \right\}$$

where $\bar{x} = \sum_{i=1}^{n} x_i/n$, $s_x^2 = \sum_{1}^{n} (x_i - \bar{x})^2/(n - 1)$, etc., has the distribution of Hotelling's Generalized Student Ratio [2]. A confidence region for a future observation (x, y) on the basis of a sample of n on a level of significance $= \alpha$ will be given by the elliptic region $T^2 \leq T_\alpha^2$ (in the x, y plane), where $T_\alpha^2 = 2(n - 1) F_0/(n - 2)$, where F_0 is the value of the F distribution (with $n_1 = 2$ and $n_2 = n - 2$ degrees of freedom) which is exceeded with probability $= 1 - \alpha$.

If P denotes the probability of a future observation falling in this ellipse, then $P = \iint_{T^2 \leq T_\alpha^2} f(x, y)\, dx\, dy$. By utilizing the fact [2] that T^2 is invariant under linear

transformations, it is not difficult to see that the distribution of P will not involve any unknown parameters, so its distribution can be calculated under the assumption $m_x = m_y = \rho = 0$, $\sigma_x = \sigma_y = 1$. Then

$$P = F(\bar{x}, \bar{y}, s_x, s_y, r) = \iint_{T^2 \le T_\alpha^2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \cdot \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}\, dx\, dy.$$

We know that $E(P) = \alpha$, and we will now calculate the variance of P by expanding P in a Taylor Series (to terms of the first order) about the point $\bar{x} = 0$, $\bar{y} = 0$, $r = 0$, $s_x = 1$, $s_y = 1$. P can clearly be put in the form

$$P = \frac{1}{2\pi} \int_{\bar{x} - T_\alpha \sqrt{\frac{n+1}{n}}\, s_x}^{\bar{x} + T_\alpha \sqrt{\frac{n+1}{n}}\, s_x} e^{-\frac{1}{2}x^2}\, dx \int_{y + rs_y\left(\frac{x-\bar{x}}{s_x}\right) - s_y \sqrt{(1-r^2)\left[T_\alpha^2\left(\frac{n+1}{n}\right) - \left(\frac{x-\bar{x}}{s_x}\right)^2\right]}}^{\bar{y} + rs_y\left(\frac{x-\bar{x}}{s_x}\right) + s_y \sqrt{(1-r^2)\left[T_\alpha^2\left(\frac{n+1}{n}\right) - \left(\frac{x-\bar{x}}{s_x}\right)^2\right]}} e^{-\frac{1}{2}y^2}\, dy$$

Taking derivatives and evaluating about the population values

$$\left[\frac{\partial P}{\partial \bar{x}}\right] = \left[\frac{\partial P}{\partial \bar{y}}\right] = \left[\frac{\partial P}{\partial r}\right] = 0,$$

$$\left[\frac{\partial P}{\partial s_x}\right] = \frac{e^{-\frac{1}{2}T_\alpha^2\left(\frac{n+1}{n}\right)}}{\pi} \int_{-T_\alpha \sqrt{\frac{n+1}{n}}}^{T_\alpha \sqrt{\frac{n+1}{n}}} \frac{x^2\, dx}{\sqrt{T_\alpha^2\left(\frac{n+1}{n}\right) - x^2}}$$

$$= \tfrac{1}{2} e^{-\frac{1}{2}T_\alpha^2\left(\frac{n+1}{n}\right)} T_\alpha^2\left(\frac{n+1}{n}\right)$$

$$\left[\frac{\partial P}{\partial s_y}\right] = \frac{e^{-\frac{1}{2}T_\alpha^2\left(\frac{n+1}{n}\right)}}{\pi} \int_{-T_\alpha \sqrt{\frac{n+1}{n}}}^{T_\alpha \sqrt{\frac{n+1}{n}}} \sqrt{T_\alpha^2\left(\frac{n+1}{n}\right) - x^2}\, dx$$

$$= e^{-\frac{1}{2}T_\alpha^2\left(\frac{n+1}{n}\right)} \cdot \tfrac{1}{2} T_\alpha^2\left(\frac{n+1}{n}\right).$$

So $$\delta P = e^{-T_\alpha^2\left(\frac{n+1}{n}\right)} \cdot \tfrac{1}{2} T_\alpha^2\left(\frac{n+1}{n}\right)[\delta s_x + \delta s_y],$$

and to terms of $0\left(\frac{1}{n}\right)$:

$$\sigma_P^2 = \frac{T_\alpha^4 e^{-T_\alpha^2}}{4n}.$$

Since for ordinary values of $\alpha(\alpha = .95$ or $.99)$ the distribution of P seems to approach normality very slowly, we will follow a suggestion of Wilks and suppose that a fairly close approximation to the distribution of P will be given by

$$\frac{\Gamma(u+v)}{\Gamma(u)\Gamma(v)} P^{u-1}(1-P)^{v-1}, \tag{1}$$

where $$u = [\alpha^2(1 - \alpha) - \alpha\sigma_P^2]/\sigma_P^2$$

$$v = [\alpha(1 - \alpha)^2 - (1 - \alpha)\sigma_P^2]/\sigma_P^2 .$$

This distribution can now be used to establish tolerance limits. For example, it follows from (1) that for a sample size $n \geq 214$, and a tolerance region given by the ellipse $T^2 = 9.21$, then $E(P) = .99$ and the Prob.$\{.985 \leq P \leq .995\} \geq .992$.

Care must be taken in the use of these and similar results, for if the distribution is not a bivariate normal one, a large error may be introduced which will not be eliminated with increasing n; however the error will probably be small when a tolerance region is found for the means $\bar{x}$, $\bar{y}$ of a future sample of k observations ($k \geq 20$) as contrasted with a tolerance region for a single observation. An exact treatment of the case when the bivariate distribution is unknown has been given by Wald in the present issue of the *Annals of Mathematical Statistics.*

REFERENCES

[1] S. S. Wilks, "Determination of sample sizes for setting tolerance limits," *Annals of Math. Stat.*, Vol. 12 (1941), pp. 91–96.

[2] Harold Hotelling, "A generalization of Student's ratio," *Annals of Math. Stat.*, Vol. 2 (1931), pp. 360–378.

A NEW APPROXIMATION TO THE LEVELS OF SIGNIFICANCE OF THE CHI-SQUARE DISTRIBUTION.

By Leo A. Aroian

Hunter College

Recent articles on the percentage points of the χ^2 distribution [1], [2], have directed my attention to a method proposed in my investigation of Fisher's z distribution [3], a method particularly useful and easily computed for n large. In addition, this method avoids interpolation. If $t = \dfrac{\chi^2 - n}{\sqrt{2n}}$, and $\alpha_3 = \sqrt{\dfrac{8}{n}}$, the measure of skewness for the χ^2 distribution, the following formulas give significance levels of t as quadratic functions of α_3, $t = a + b\alpha_3 + c\alpha_3^2$. The values of a, b, and c were found by the usual method of least squares, fitting each formula to the values of t [4] for $\alpha_3 = 0$, ± 0.1, ± 0.2, ± 0.3, and ± 0.4. Then the value of a in each instance was adjusted to give the proper value of t when $\alpha_3 = 0$: e.g. the constant term by the method of least squares for the 1 per cent point is 2.32633 which we change to 2.32635. The range $|\alpha_3| \leqq .4$ corresponds to $n \geqq 50$, but the formulas are quite satisfactory for $n \geqq 30$. Formulas for t when $|\alpha_3| > .4$ [3] are easily derived, but such results while more accurate in the range

TABLE I[1]

n \ P	.995	.99	.95	.90	.75	.50	.25	.10	.05	.01	.005
30	13.744	14.925	18.491	20.604	24.486	29.338	34.793	40.246	43.767	50.914	53.713
	13.7997	14.9649	18.4960	20.6004	24.4764	29.3346	34.7987	40.2559	43.7754	50.9015	53.6883
	13.7867	14.9535	18.4926	20.5992	24.4776	29.3360	34.7998	40.2560	43.7729	50.8922	53.6720
40	20.669	22.139	26.5080	29.055	33.668	39.337	45.610	51.796	55.753	63.710	66.802
	20.7121	22.1703	26.5114	29.0514	33.6597	39.3346	45.6155	51.8048	55.7600	63.6961	66.776
	20.7065	22.1643	26.5093	29.0505	33.6603	39.3354	45.6160	51.8050	55.7585	63.6907	66.7659
50	27.957	29.685	34.7634	37.693	42.949	49.336	56.328	63.159	67.501	76.172	79.523
	27.9920	29.7096	34.7656	37.6894	42.9418	49.3346	56.3333	63.1669	67.5057	76.1568	79.496
	27.9907	29.7067	34.7642	37.6886	42.9421	49.3349	56.3336	63.1671	67.5048	76.1539	79.4900
75	47.178	49.457	56.0538	59.799	66.422	74.335	82.853	91.055	96.2135	106.408	110.313
	47.2021	49.4741	56.0546	59.7951	66.4169	74.3346	82.8581	91.0611	96.2168	106.392	110.286
	47.2059	49.4748	56.0540	59.7944	66.4167	74.3343	82.8582	91.062	96.2160	106.393	110.286
100	67.303	70.049	77.9294	82.362	90.138	99.335	109.137	118.493	124.340	135.820	140.193
	67.3210	70.0620	77.9296	82.3586	90.1336	99.3346	109.1416	118.498	124.342	135.804	140.166
	67.3276	70.0648	77.9295	82.3581	90.1332	99.3341	109.141	118.498	124.342	135.807	140.169

[1] First value in cell by method of Wilson and Hilferty.
Second value in cell by (1) or (2).
Third value is correct result.

TABLE II[2]

P	.9999	.999	.975	.80	.70	.60	.40	.30	.20	.025	.001	.0001
$n = 30$	9.33	11.62	16.7962	23.3631	25.5064	27.4402	31.3144	33.5290	36.2494	46.9821	59.73	67.68
	9.226	11.568	16.7908	23.364	25.508	27.4436	31.3183	33.530	36.250	46.9792	59.683	67.590

[2] First value by (1) or (2).
Second value correct result.

$30 \leqq n < 50$ would be considerably less accurate in the region $n \geqq 50$. After t is calculated, $\chi^2 = n + \sqrt{2n}\,t$. The formulas are:

(1)
$$t_{50\%} = -.16636\alpha_3$$
$$t_{40\%} = .25335 - .15567\alpha_3 - .012276\alpha_3^2$$
$$t_{30\%} = .52440 - .12058\alpha_3 - .024541\alpha_3^2$$
$$t_{25\%} = .67449 - .090613\alpha_3 - .030693\alpha_3^2$$
$$t_{20\%} = .84162 - .048433\alpha_3 - .036788\alpha_3^2$$
$$t_{10\%} = 1.28155 + .107033\alpha_3 - .04797\alpha_3^2$$
$$t_{5\%} = 1.64485 + .28392\alpha_3 - .04902\alpha_3^2$$
$$t_{2.5\%} = 1.95997 + .47228\alpha_3 - .04304\alpha_3^2$$
$$t_{1\%} = 2.32635 + .73330\alpha_3 - .024957\alpha_3^2$$
$$t_{.5\%} = 2.5758 + .93600\alpha_3 - .00377\alpha_3^2$$
$$t_{.1\%} = 3.0903 + 1.4190\alpha_3 + .05667\alpha_3^2$$
$$t_{.01\%} = 3.7200 + 2.1260\alpha_3 + .17449\alpha_3^2$$

The maximum error for t in the range $|\alpha_3| \leqq .4$, is 2 in the fourth significant figure, 1 in the fourth significant figure, 6 in fifth, 3 in fifth, 3 in fifth, 1 in fifth, 1 in fifth, 3 in fifth, 4 in fifth, 4 in fifth, 4 in fifth and 4 in fourth significant figures respectively for the .01%, .1%, .5%, 1%, 2.5%, 5%, 10%, 20%, 25%, 30%, 40%, and 50% points respectively. The error increases outside the indicated range. In addition

(2)
$$t_{99.99\%} = -3.7200 + 2.1260\alpha_3 - .17449\alpha_3^2$$
$$t_{99.9\%} = -3.0903 + 1.4190\alpha_3 - .05667\alpha_3^2$$

and similarly for other percentage points. These are obtained from (1) by replacing α_3 by $-\alpha_3$ and t by $-t$.

We compare results obtained by these methods against those of Wilson and Hilferty [2]. In all cases except at the 95% level the method here proposed is superior. Table I compares the two methods. It was copied from [2] except for the corrections in the Wilson and Hilferty method for the 95% level and in the accurate value for χ^2 at the 5% level for $n = 75$, 96.2160 in place of 96. 11. Table II gives comparisons for other levels when $n = 30$.

REFERENCES

[1] C. M. Thompson, "Table of percentage points of the χ^2 distribution," *Biometrika*, Vol. 32, Part 2.

[2] M. Merrington, "Numerical approximations to the percentage points of the χ^2 distribution," *Biometrika*, Vol. 32, Part 2.

[3] L. A. Aroian, "A study of R. A. Fisher's z distribution and the related F distribution," *Annals of Math. Stat.*, Vol. 12 (1941).

[4] L. R. Salvosa, "Tables of Pearson's Type III function," *Annals of Math. Stat.*, Vol. 1.

Second value correct result.

NEWS AND NOTICES

Readers are invited to submit to the Secretary of the Institute news items of general interest

Personal Items

Associate Professor H. P. Evans of the Mathematics Department of the University of Wisconsin has been promoted to a professorship.

Assistant Professor Willy Feller of the Mathematics Department of Brown University has been promoted to an associate professorship.

Dr. Carl F. Kossack of the Mathematics Department of the University of Oregon has been promoted to an assistant professorship.

Dr. Eugene Lukacs has been appointed to an assistant professorship in the Mathematics Department of Illinois College.

Professor E. B. Mode has been made chairman of the Mathematics Department of Boston University.

Mr. Charles R. Mummery has been made Product Quality Engineer at the Scioto Ordinance Plant of the U. S. Rubber Company.

Professor H. L. Rietz has retired after twenty-five years of service as Head of the Mathematics Department of the University of Iowa.

The Foundation for the Study of Cycles has announced that a medal will be awarded to the individual making the most significant contribution to cycle research during 1943. Communications should be addressed to: Professor Ellsworth Huntington, Yale University, New Haven, Connecticut.

Obituary

Professor Edward L. Dodd of the Mathematics Department of the University of Texas died on January 9, 1943 at the age of sixty-seven years. He was a charter member of the Institute. He was elected as one of the Vice-Presidents of the Institute for 1943. His contributions to mathematical statistics consist of numerous research papers on probability; on general mean functions of statistical variables; and on statistical theory of periodicities.

Stanford Courses in Statistical Methods of Quality Control

A novel procedure in adult education, and particularly in statistical education, took place last summer at Stanford University, when courses in the Shewhart statistical methods of quality control were offered in short intensive courses. There were two courses, one on the campus at Stanford University, July 17–26, and the other in Los Angeles, September 20–27. The first course covered ten full days, and the second eight. Both courses ran eight hours per day, Saturdays and Sundays included.

The features of the course may be described by the following points:

1. The courses were short, thus making it possible for men in industry to attend.
2. The number of hours' instruction was sufficient to cover the field adequately.
3. The instruction covered a wide range of points of view.
4. The students were picked delegates sent by industry.
5. The courses are being followed up with monthly meetings in Los Angeles and San Francisco.

Intensive courses of this character were first suggested by Dr. W. Edwards Deming in April of 1942, while he was temporarily detailed to the office of the Chief of Ordnance in the War Department, and the first course actually commenced just three months later. By giving the course to men already in industry, the yield obtained was manyfold higher than can be expected from a regular college course. Contributions and reports made by the delegates subsequent to the courses supply abundant foundation for this statement.

West Coast industry and the Army and Navy ordnance districts sent 32 delegates to the West Coast, and 31 to the second. Through the efforts of Professor Eugene Grant of Stanford, industry and the Army and Navy were persuaded to send some of their most valued officials. The instruction was organized by Professor Holbrook Working. Both he and Professor Grant took an active part in the instruction, which was supplemented in both courses by Dr. W. Edwards Deming as an exponent of government and industrial sampling. In the first course, Mr. Charles R. Mummery of the Hoover Company served as an instructor from the viewpoint of industry. In the second course (the one in Los Angeles), Mr. Ralph E. Wareham of the General Electric Company occupied the industrial corner of the square of instruction. The expense of the instructors was paid out of ESMWT funds (Office of Education). Montly follow-up courses in San Francisco and Los Angeles, under the direction of Professors Working and Grant, supply the necessary power for maintaining momentum, and for gathering the men together for directed study and consultation.

The demand for men trained in this line far exceeds the supply, and there are movements afoot to provide similar courses in a number of industrial cities. Three-day courses in a dozen or more key ordnance cities were held last fall by the Ordnance Department. The lecturers were Messrs. G. D. Edwards and Harold F. Dodge of the Bell Telephone Laboratories, and Mr. G. Rupert Gause of the Aberdeen Proving Ground, now with the Army Ordnance in Washington. These courses and the Stanford courses alleviated the situation considerably, but further instruction is needed.

Junior Membership in the Institute

At the annual election for 1942 which was held by mail ballot because of the postponement of the Annual Meeting, constitutional amendments were approved which created a new grade of membership in the Institute, known as Junior

Membership. It is hoped that this provision for Junior Membership will stimulate interest in mathematical statistics at the advanced undergraduate level in colleges and universities.

The Board of Directors have approved the following rules governing Junior Membership:

1. Any undergraduate student of a collegiate institution is eligible for election as a Junior member of the Institute of Mathematical Statistics provided that he or she is sponsored by a member of the Institute.
2. The annual dues ($2.50) must be submited with the application.
3. Annual membership shall coincide with the calendar year and the Junior Member shall receive a complete volume of the *Annals of Mathematical Statistics* for the year in which he or she is elected.
4. Junior Membership shall be limited to a term of two years, but a Junior Member may apply for transfer to ordinary membership at the beginning of his second year.

For the convenience of any Institute member who may wish to sponsor a Junior Member an application blank is provided at the back of this issue of the *Annals*. Additional blanks may be obtained from the Secretary of the Institute.

Announcement of May Meeting in New York

There will be a joint meeting between the Institute and the American Society of Mechanical Engineers on Saturday, May 29, 1943, at the Engineering Societies Building, 29 West 39th Street, N. Y.

The meeting will consist of two sessions on industrial applications of mathematical statistics. The topics are as follows:

Morning Session, 10 A.M.

Chairman: Harold Hotelling

1. J. Wolfowitz, *On the Theory of Runs with some Applications to Quality Control.*
2. Churchill Eisenhart, *On the Presentation of Data as Evidence.*

Afternoon Session, 2 P.M.

Chairman: W. A. Shewhart

1. H. F. Dodge, *A Sampling Inspection Plan for Continuous Production.*
2. L. C. Young, *Tolerances and Product Acceptability.*

ANNUAL REPORT OF THE PRESIDENT OF THE INSTITUTE

Ordinarily at the business meeting and at the luncheon customarily held as part of the annual meeting of the Institute, the President has the opportunity to make public acknowledgement to those individuals, aside from the officers, who served the Institute during the year, and to have his say concerning past progress and future plans. This year it appears that the pages of the *Annals* must be used for this purpose.

As the result of proposals made and approved at our last regular annual meeting in New York City, a larger number of special committees than usual were appointed for 1942 and thus more members than before were specifically asked to participate in the affairs of the Institute. The Institute is much indebted to these individuals for the way in which they responded.

Professors A. T. Craig, Harold Hotelling, and S. S. Wilks, Chairman, constituted a committee to study the Board of Directors of the Institute and the formal connection between the Institute and its journal, *The Annals of Mathematical Statistics*. Their recommendations were incorporated in amendments to the constitution and by-laws recently approved by the Institute. The Board was increased in size and given greater continuity by including in it the two previous presidents, and the editor of the *Annals*, ex officio, and by increasing the term of the Secretary-Treasurer to three years.

The new class of junior memberships is the result of a study of this question by a committee composed of Professors J. H. Bushey, Boyd Harshbarger, and G. W. Snedecor, Chairman. Regulations, since approved by the Board, under which local chapters of the Institute may be formed, were drawn up by a committee consisting of Dr. C. F. Kossack, Professor H. D. Larsen, and Professor B. H. Camp, Chairman.

Dr. L. A. Aroian, Dr. J. F. Daly, Mr. H. F. Dodge, and Professor W. D. Baten, Chairman, as a committee agreed to assist the Institute by endeavoring to bring the *Annals of Mathematical Statistics* to the favorable attention of municipal, industrial, and college libraries which had not been subscribers to it. As a result of their fine work, a good number of domestic libraries has been added to our subscription list, thus serving to counterbalance our losses abroad.

The Program Committees for the year consisted of Professor Churchill Eisenhart and Mr. E. C. Molina, Chairman, for the September meeting in Poughkeepsie, New York, and of Professor P. S. Dwyer and Dr. W. E. Deming, chairman, for the projected Cleveland meeting. The Institute is always much indebted to those who do the work of arranging its programs for meetings, but this year we owe Dr. Deming a special acknowledgement, who prepared an excellent program for Cleveland, then one for a New York meeting under extremely short notice when the Cleveland meeting was cancelled, and then had that meeting also cancelled. Dr. W. R. Van Voorhis acted as our representative on the Committee on Local Arrangements for the meeting planned for Cleveland.

The membership committee appointed for 1942 was made up of Dr. W. E. Deming, Professor E. L. Dodd, and Professor A. T. Craig, chairman. After Professor Craig took up his commission in the Navy, Professor B. H. Camp agreed to take his place on this committee.

For some years Professor A. T. Craig has generously acted as custodian of our files of back numbers of the *Annals*. This service to the Institute has been taken over by Professor L. A. Knowler, who has been of much assistance. Dr. W. E. Blanche did a considerable amount of work in connection with finding advertisers for the *Annals*.

Our annual meetings had been increasingly successful in recent years and it was a real sacrifice for the Institute to forego the one planned for 1942. Though the war is demonstrating in still more ways and places the importance of sound statistical methods, for the present it imposes serious responsibilities on the friends of the Institute. A reading of the report of our faithful and efficient Secretary-Treasurer will amplify this statement. Of the present Board, Professors Olds, Wilks and Craig met in Pittsburgh January 23 and 24 to consider some of our problems. Though there seems no prospect of a national meeting in the coming year, it is hoped that some local meetings can be held and that in other ways we can keep up the activities of the Institute. In particular there exists the opportunity of organizing local chapters of the Institute in the larger centers which would be particularly valuable now. In industrial areas we may contribute to the war effort as well as promote an important aspect of mathematical statistics by endeavoring to be useful in the development and application of industrial statistics. It is clear that the Institute needs the loyal support of its membership now as much as ever before if it is to fulfill the functions for which it was founded.

CECIL C. CRAIG,
President.

December 31, 1942.

ANNUAL REPORT OF THE SECRETARY-TREASURER OF THE INSTITUTE

On September 8–9 the Institute met at Vassar College in conjunction with the American Mathematical Society and the Mathematical Association of America. Mr. E. C. Molina and Professor Churchill Eisenhart were in charge of the program. Fifty eight members of the Institute attended the meeting. The Annual Meeting, originally scheduled for Cleveland then transferred to New York City, was finally postponed at the request of the Office of Defense Transportation. At the present time it seems that this meeting will have to be abandoned entirely and the Institute must be content with holding local meetings in some of the larger cities.

Because of the postponement of the Annual Meeting, the annual election was held by mail. The following officers were elected: Professor Cecil C. Craig, President; Professors Edward L. Dodd and Abraham Wald, Vice-Presidents; and Professor Edwin G. Olds, Secretary-Treasurer. Nine amendments to the Constitution and six amendments to the By-Laws were proposed and accepted by a two-thirds majority of those voting. Professor K. L. Fetters acted as teller.

During the past year the Secretary has cooperated with industrial concerns and government agencies in locating statistically trained personnel to fill positions created by the emergency. Members of the Institute are requested to keep the Secretary informed regarding the availability of such personnel.

The death of one member of the Institute has been reported since the last

annual meeting—Dr. Robert Henderson, former Vice-President and Actuary of the Equitable Life Assurance Society.

The following financial statement covers the period from December 10, 1941 to December 10, 1942 (the books and records of the Treasurer have been audited by Mr. George E. Niver and found to be in agreement with the statement as submitted):

FINANCIAL STATEMENT

December 10, 1941, to December 10, 1942

Receipts

Balance on Hand, December 10, 1941		$1,561.54
Dues		2,293.15
Subscriptions		1,289.74
Sales of Back Numbers		1,393.65
Cumulative Index		5.00
Miscellaneous		28.50
Total Receipts		$6,571.58

Expenditures

Annals Office		
Editorial Expenses		$127.98
Waverly Press		
Printing and Mailing Annals—4 issues		3,227.34
Back Numbers Office		
Purchase of back numbers from H. C. Carver	$355.77	
Reprinting 300 copies of Vol. V, No. 4	142.16	
		$497.93
Library Committee		16.77
Secretary-Treasurer's Office		
Printing and Supplies	$53.58	
Binding	30.00	
Postage	139.16	
Clerical Help	215.50	
		$438.24
Printing Programs for Meetings		101.11
Miscellaneous		7.08
Total Expenditures		$4,416.45
Balance on Hand, December 10, 1942		2,155.13
		$6,571.58

In comparison with the financial condition of the Institute at the end of 1941, the receipts from dues, subscriptions, and sales of back numbers have increased more than $800. This is mostly due to a large increase in the sales of back numbers and a net increase of fifty members. The increase in expenditures of the Institute was accounted for by the increased cost in printing the *Annals*. This marks the beginning of a trend which seems likely to continue throughout the war.

It would seem over-optimistic to expect that the financial situation of the Institute would continue to show marked improvement in 1943. Present indications suggest that we shall be very fortunate to avoid a considerable deficit in operations. A large number of our foreign subscribers have not renewed and we face considerable difficulty in delivery of the *Annals* to those still in force. The large increase in the sales of back numbers was due to a rather successful effort to persuade domestic libraries to provide themselves with complete sets of back numbers while the issues were still available. The Institute faces an increase in operating expenses and an advance in the cost of producing the *Annals*. The full cooperation of all members is needed if we are to avoid a decrease in the work of the Institute during 1943.

EDWIN G. OLDS,
Secretary-Treasurer.

December 31, 1942.

On behalf of the Board of Directors of the Institute, I regret to announce the sudden death of Vice-President E. L. Dodd, on January 9, 1943, shortly after this report was written. Dr. W. E. Deming was appointed by the Board of Directors to fill the vacancy created by Vice-President Dodd's death.

E. G. O.

CONSTITUTION
OF THE
INSTITUTE OF MATHEMATICAL STATISTICS

ARTICLE I

NAME AND PURPOSE

1. This organization shall be known as the Institute of Mathematical Statistics.
2. Its object shall be to promote the interests of mathematical statistics.

ARTICLE II

MEMBERSHIP

1. The membership of the Institute shall consist of Members, Junior Members, Fellows, Honorary Members, and Sustaining Members.
2. Voting members of the Institute shall be (a) the Fellows, and (b) all others, Junior Members excepted, who have been members for twenty-three months prior to the date of voting.
3. No person shall be a Junior Member of the Institute for more than a limited term as determined by the Committee on Membership and approved by the Board of Directors.

ARTICLE III

OFFICERS, BOARD OF DIRECTORS, AND COMMITTEE ON MEMBERSHIP

1. The Officers of the Institute shall be a President, two Vice-Presidents, and a Secretary-Treasurer. The terms of office of the President and Vice-Presidents shall be one year

and that of the Secretary-Treasurer three years. Elections shall be by majority ballots at Annual Meetings of the Institute. Voting may be in person or by mail.

(a) Exception. The first group of Officers shall be elected by a majority vote of the individuals present at the organization meeting, and shall serve until December 31, 1936.

2. The Board of Directors of the Institute shall consist of the Officers, the two previous Presidents, and the Editor of the Official Journal of the Institute.

3. The Institute shall have a Committee on Membership composed of three Fellows. At their first meeting subsequent to the adoption of this Constitution, the Board of Directors shall elect three members as Fellows to serve as the Committee on Membership, one member of the Committee for a term of one year, another for a term of two years, and another for a term of three years. Thereafter the Board of Directors shall elect from among the Fellows one member annually at their first meeting after their election for a term of three years. The president shall designate one of the Vice-Presidents as Chairman of this Committee.

ARTICLE IV

Meetings

1. A meeting for the presentation and discussion of papers, for the election of Officers, and for the transaction of other business of the Institute shall be held annually at such time as the Board of Directors may designate. Additional meetings may be called from time to time by the Board of Directors and shall be called at any time by the President upon written request from ten Fellows. Notice of the time and place of meeting shall be given to the membership by the Secretary-Treasurer at least thirty days prior to the date set for the meeting. All meetings except executive sessions shall be open to the public. Only papers accepted by a Program Committee appointed by the President may be presented to the Institute.

2. The Board of Directors shall hold a meeting immediately after their election and again immediately before the expiration of their term. Other meetings of the Board may be held from time to time at the call of the President or any two members of the Board. Notice of each meeting of the Board, other than the two regular meetings, together with a statement of the business to be brought before the meeting, must be given to the members of the Board by the Secretary-Treasurer at least five days prior to the date set therefor. Should other business be passed upon, any member of the Board shall have the right to reopen the question at the next meeting.

3. The Committee on Membership shall hold a meeting immediately after the annual meeting of the Institute. Further meetings of the Committee may be held from time to time at the call of the Chairman or any member of the Committee provided notice of such call and the purpose of the meeting is given to the members of the Committee by the Secretary-Treasurer at least five days before the date set therefor. Should other business be passed upon, any member of the Committee shall have the right to reopen the question at the next meeting.

4. At a regularly convened meeting of the Board of Directors, four members shall constitute a quorum. At a regularly convened meeting of the Committee on Membership, two members shall constitute a quorum.

ARTICLE V

Publications

1. The *Annals of Mathematical Statistics* shall be the Official Journal for the Institute. The Editor of the *Annals of Mathematical Statistics* shall be a Fellow appointed by the

Board of Directors of the Institute. The term of office of the Editor may be terminated at the discretion of the Board of Directors.

2. Other publications may be originated by the Board of Directors as occasion arises.

ARTICLE VI

Expulsion or Suspension

1. Except for non-payment of dues, no one shall be expelled or suspended except by action of the Board of Directors with not more than one negative vote.

ARTICLE VII

Amendments

1. This constitution may be amended by an affirmative two-thirds vote at any regularly convened meeting of the Institute provided notice of such proposed amendment shall have been sent to each voting member by the Secretary-Treasurer at least thirty days before the date of the meeting at which the proposal is to be acted upon. Voting may be in person or by mail.

BY-LAWS

ARTICLE I

Duties of the Officers, the Editor, Board of Directors, and Committee on Membership

1. The President, or in his absence, one of the Vice-Presidents, or in the absence of the President and both Vice-Presidents, a Fellow selected by vote of the Fellows present, shall preside at the meetings of the Institute and of the Board of Directors. At meetings of the Institute, the presiding officer shall vote only in the case of a tie, but at meetings of the Board of Directors he may vote in all cases. At least three months before the date of the annual meeting, the President shall appoint a Nominating Committee of three members. It shall be the duty of the Nominating Committee to make nominations for Officers to be elected at the annual meeting and the Secretary-Treasurer shall notify all voting members at least thirty days before the annual meeting. Additional nominations may be submitted in writing, if signed by at least ten Fellows of the Institute, up to the time of the meeting.

2. The Secretary-Treasurer shall keep a full and accurate record of the proceedings at the meetings of the Institute and of the Board of Directors, send out calls for said meetings and, with the approval of the President and the Board, carry on the correspondence of the Institute. Subject to the direction of the Board, he shall have charge of the archives and other tangible and intangible property of the Institute, and once a year he shall publish in the *Annals of Mathematical Statistics* a classified list of all Members and Fellows of the Institute. He shall send out calls for annual dues and acknowledge receipt of same; pay all bills approved by the President for expenditures authorized by the Board or the Institute; keep a detailed account of all receipts and expenditures, prepare a financial statement at the end of each year and present an abstract of the same at the annual meeting of the Institute after it has been audited by a Member or Fellow of the Institute appointed by the President as Auditor. The Auditor shall report to the President.

3. Subject to the direction of the Board, the Editor shall be charged with the responsibility for all editorial matters concerning the editing of the *Annals of Mathematical Statistics*. He shall, with the advice and consent of the Board, appoint an Editorial Commit-

tee of not less than twelve members to co-operate with him; four for a period of five years, four for a period of three years, and the remaining members for a period of two years, appointments to be made annually as needed. All appointments to the Editorial Committee shall terminate with the appointment of a new Editor. The Editor shall serve as editorial adviser in the publication of all scientific monographs and pamphlets authorized by the Board.

4. The Board of Directors shall have charge of the funds and of the affairs of the Institute, with the exception of those affairs specifically assigned to the President or to the Committee on Membership. The Board shall have authority to fill all vacancies ad interim, occurring among the Officers, Board of Directors, or in any of the Committees. The Board may appoint such other committees as may be required from time to time to carry on the affairs of the Institute.

5. The Committee on Membership shall prepare and make available through the Secretary-Treasurer an announcement indicating the qualifications requisite for the different grades of membership.

ARTICLE II

Dues

Members shall pay five dollars at the time of admission to membership and shall receive the full current volume of the Official Journal. Thereafter, Members shall pay five dollars annual dues. The annual dues of Junior Members shall be two dollars and fifty cents. The annual dues of Fellows shall be five dollars. The annual dues of Sustaining Members shall be fifty dollars. Honorary Members shall be exempt from all dues.

(a) Exception. In the case that two Members of the Institute are husband and wife and they elect to receive between them only one copy of the Official Journal, the annual dues of each shall be three dollars and seventy-five cents.

2. Annual dues shall be payable on the first day of January of each year.

3. The annual dues of a Fellow, Member, or Junior Member include a subscription to the Official Journal. The annual dues of a Sustaining Member include two subscriptions to the Official Journal.

4. It shall be the duty of the Secretary-Treasurer to notify by mail anyone whose dues may be six months in arrears, and to accompany such notice by a copy of this Article. If such person fail to pay such dues within three months from the date of mailing such notice, the Secretary-Treasurer shall report the delinquent one to the Board of Directors, by whom the person's name may be stricken from the rolls and all privileges of membership withdrawn. Such person may, however, be re-instated by the Board of Directors upon payment of the arrears of dues.

ARTICLE III

Salaries

1. The Institute shall not pay a salary to any Officer, Director, or member of any committee.

ARTICLE IV

Amendments

1. These By-Laws may be amended in the same manner as the Constitution or by a majority vote at any regularly convened meeting of the Institute, if the proposed amendment has been previously approved by the Board of Directors.